PERIODICAL ROOM
GENERAL LIBRARY
UNIV. OF MICH.

DEC 5 - 1932

VOL. XXIII NOVEMBER, 1932 NO. 8

The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching

CONTENTS

How to Construct a Nomogram 561
HAROLD D. GRIFFIN

Interpretation of Correlation on the Basis of Common Elements . . 578
P. H. NYGAARD

Reliability of Group Distinctions 586
RAYMOND FRANZEN AND MAHEW DERRYBERRY

Rating Tendencies of Personally Selected Judges 594
RICHARD STEPHEN UHRBROCK

A Standardized Test of Prose Appreciation for Junior High School Pupils 604
HERBERT A. CARROLL

The Constancy of the IQ's of Gifted Children 607
CLAUDE L. NEMZEK

A Biological View of Behavior Modification 611
J. STANLEY GRAY

The Curve of Forgetting: Its Statistical Application 621
R. J. WHERRY

The Rôle of Affective Tone in Learning to Spell 625
ARDEN FRANDSEN

The Relative Efficiency of the Whole, Part and Progressive Part Methods When Trials Are Massed—A Minor Experiment . . . 632
J. B. STROUD AND C. W. RIDGEWAY

Book Reviews 635

$5.00 per Year • Published Monthly Except June to August by

WARWICK & YORK, INC.

BALTIMORE, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879; additional entry as second class matter at York, Pa.

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

ESTABLISHED 1910

Board of Editors

HAROLD RUGG

J. CARLETON BELL ARTHUR I. GATES RUDOLF PINTNER E. L. THORNDIKE
FRANK N. FREEMAN V. A. C. HENMON L. M. TERMAN GOODWIN B. WATSON

The price of the Journal is $5.00 a year in the United States; $5.30 in foreign countries. Subscriptions for less than a year are charged at the rate of 70 cents for each number ordered. Back volumes are $6.00 each; back issues, 90 cents each except when three or more years old, and then $1.00 each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which the change is to take effect. Claims for non-receipt of an issue will not be honored unless made within two weeks after receipt of the next succeeding number.

Manuscripts for publication, books or other materials for review, and correspondence should be addressed to the publishers.

WARWICK AND YORK, INC., *Baltimore, Maryland*

TEACHERS' HEADQUARTERS

European Plan

Entirely Fire-Proof

Centrally Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Psychology of Learning
1928 Revision By WILLIAM H. PYLE
$2.20+12¢ postage
Warwick & York, Inc., *Baltimore, Maryland*

The Acid Test

What more exacting test can there be of dictionary leadership than the judgment of the educational world?

Every State that has adopted an unabridged dictionary for use in its schools has chosen exclusively

WEBSTER'S NEW INTERNATIONAL DICTIONARY

The Colleges voted overwhelmingly in favor of Webster as the standard of pronunciation, in answer to questions submitted by the Chicago Woman's Club.

Nearly 100% of all schoolbooks are based on the New International for pronunciation, spelling, compounding, and division of words.

The New International has been universally accepted by those best fitted to judge.

It is in accord with the best modern practice.

Use it in your own schools.

Write for our free booklets of interesting ways to teach the use of the dictionary.

GET THE BEST

G. & C. MERRIAM COMPANY
SPRINGFIELD, MASS.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XXIII November, 1932 Number 8

HOW TO CONSTRUCT A NOMOGRAM

HAROLD D. GRIFFIN

Nebraska State Teachers College, Wayne, Nebraska

DEFINITIONS

A nomogram (from the Greek *nomos*, law + *gramma*, writing) is a graphic representation of a formula, so plotted that the variables appear on scales, their relative magnitudes and relative positions being such that *corresponding values* of the variables are found on a *straight line* crossing the scales. For this reason a nomogram is often called an alinement chart or a straight-line diagram. By *corresponding values* we mean any set of values which satisfy the equation represented by the nomogram. Therefore, the crossing line is termed an *isopleth*, from the Greek *isos*, equal + *plethos*, quantity, number, magnitude. A satisfactory isopleth is made by scratching a straight line with a needle on the underside of a strip of good weight celluloid.

Nomograms are far simpler to read than *abacs*, or Cartesian diagrams, that is, formulas plotted on coordinate and logarithmic papers. Abacs have a distressing confusion of lines which are hard to follow and difficult to interpolate. Compare the abac for the probable error of a correlation coefficient on p. 19 of Pearson's tables* with the same formula charted as a nomogram in this article.

ORIGIN OF NOMOGRAMS

Maurice d'Ocagne† invented nomograms about 1884. His first book on the subject was *Nomographie* (1891). This and his *Traite de*

* Pearson, Karl: "Tables for Statisticians and Biometricians." Part I, second Edition. Cambridge University Press, London, 1924.

A scholarly complilation; excellent for reference, but not espically necessary for the layman.

† Maurice d'Ocagne: French scholar, mathematician and engineer. For an account of his life and works consult "Qui etes-vous? Annuaire des contemporains

nomographie (1899, 1921) constitute two of the classics of the science. The literature is voluminous, as may be judged from the annotated bibliography with this article.

PRINCIPLE OF THE NOMOGRAM

The principle of the nomogram may be illustrated by a simple experiment. Mark off on the edge of a twelve-inch strip of cardboard the graduations from a foot ruler, but label every half-inch consecutively, so that the major marks will be from 0 to 24 rather than from 0 to 12. Now lay two twelve-inch rulers on the table about twelve inches apart, with their zero ends at the edge of the table, and their axes parallel and at right angles to the table edge. Put the cardboard ruler just halfway between the other two, with its zero end at the edge of the table. Hold a piece of thread tightly across these three rulers—say from the 4 on the first ruler to the 10 on the third ruler. If the rulers are strictly parallel, with the second one centrally located, the thread will cut it at 14. $4 + 10 = 14$. Stretch the thread from the 8 on the first ruler to the 2 on the third. The thread will cut the second at 10. $8 + 2 = 10$. It is apparent that we have here a graphical device for performing additions. As such, its usefulness would be very much limited. But suppose we should replace our uniform scales by logarithmic scales. Now, instead of multiplying two natural numbers, we may add their logarithms.* Therefore, the nomogram may be employed for multiplication by constructing its axes with logarithmic readings. Let us reverse one of our outside scales and experiment with our thread. We find that we have a subtracting device. Substitute logarithmic scales and we may perform divisions.

A study of the three rulers reveals, then, two things: First, that the parallel-axis nomogram is essentially a device for addition and subtraction, and that its value as a multiplying and dividing instrument is due to the peculiar properties of logarithms. Second, that the middle axis, its graduations and its position, seems to bear a proportional relationship to the other two axes. We find that this depends on a simple and familiar property of similar triangles.

---

notices biographiques." G. Ruffy, Paris, yearly; and "Enciclopedia universal ilustrada." Vol. XXXIX, pp. 483–484. Hijos de J. Espasa, Barcelona.

* This principle was known as early as 1620. Cajori, F.: The History of Gunter's Scale and the Slide-rule during the Seventeenth Century. *University of California Publications in Mathematics*, Vol. I, 1920, pp. 187–209.

GEOMETRIC THEORY OF THE NOMOGRAM

Let X, Y, and Z (Chart I) be three parallel axes, with *abc* the base line. Now cut these lines by an index line *def* and examine the relationships between the three distances *ad*, *be*, and *ef*, which distances

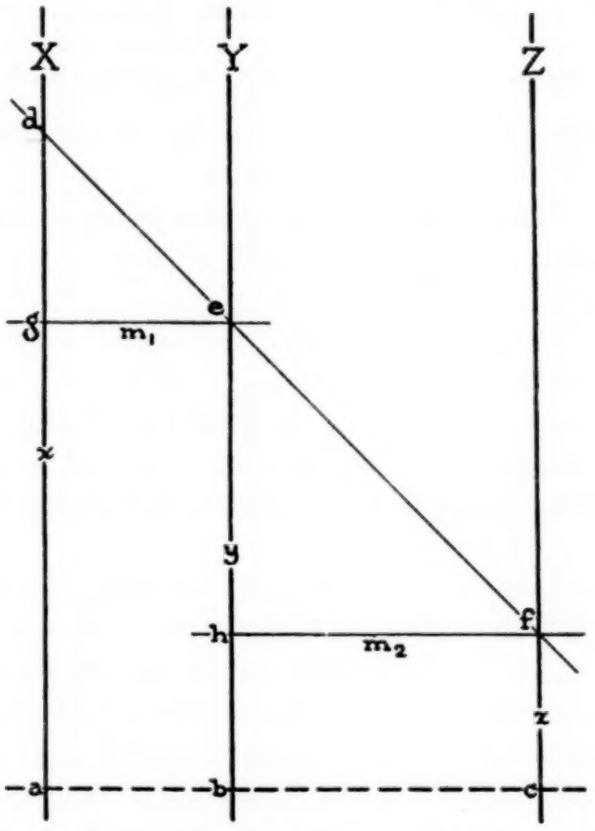

CHART I.—Illustrating the geometry of the nomogram.

we shall name respectively x, y, and z. If we draw m_1* from e and m_2 from f, parallel to the base line, then the triangles formed (*dge* and *ehf*) will be similar, with $gd/he = m_1/m_2$ which, as $ag = y$ and $bh = z$, is $(x - y)/(y - z) = m_1/m_2$ or $m_2(x - y) = m_1(y - z)$. Therefore, $m_2x + m_1z = (m_1 + m_2)y$, or, $x/m_1 + z/m_2 = y/m_1m_2/(m_1 + m_2)$, which indicates that the graduations on axis y are determined by $m_1m_2/(m_1 + m_2)$.

* The terms m_1 and m_2 in this connection will be considered as ratios. In reality they are the symbols for the *scale units*, technically called *moduli* (singular, *modulus*), on x and z respectively; but, from the properties of similar triangles, they serve also to mark the distance that Y will be from X and Z.

This is the geometry of Chart I. From it we see that if the number of graduations in y is to be equal to the sum of the graduations in x and z, the graduations on axis Y must be smaller than graduations representing the same values on either X or Z by a multiple of two or more, depending on the relation of m_1 to m_2. Therefore, when Y is halfway between X and Z this multiple will be 2, and the Y-axis graduations will be half the length of the X- and Z-axis graduations.

STEPS IN CONSTRUCTING A NOMOGRAM

In constructing Chart II, "Given deviation (d) and frequency (f), to find fd and fd^2," we proceed as follows.

Formulas.—1. There are two formulas here, $fd = f \times d$, and $fd^2 = f \times d^2$. These may be put in addition form as $\log (fd) = \log d + \log f$, and $\log (fd^2) = 2 \log d + \log f$.

Preliminary Calculations.—2. Fix limits for the values of these variables—1 to 25 for the deviations, 1 to 100 for the frequencies, and, by calculating the limits, 1 to 2500 for fd and 1 to 62,500 for fd^2. The logarithmic difference between 1 and 25 is 1.3979, between 1 and 100 is 2.0000, between 1 and 2500 is 3.3979, and between 1 and 62,500 is 4.7959.

Plotting the Axes.—3. Draw two parallel axes, X and Z, ten inches apart and fifteen inches long. On these we shall construct the logarithmic scales x and z corresponding to the deviations and frequencies respectively. Now, as the logarithmic difference between the 1 and the 25 for the deviations on the x-scale is 1.3979, and as the length of the axis is fifteen inches, the modulus—that is, the unit distance, which in logarithms is from arithmetic figures 1 to 10—will be 15/1.3979, or about 10.73 inches. Therefore, $m_1 = 10.73$ inches. The modulus for the z-scale will be $15/2 = 7.5$ inches. Therefore, $m_2 = 7.5$ inches.

4. The Y-axis will then be erected parallel to the X- and Z-axes, and at the distance $10.73/7.5 = x/(10 - x)$, or 5.886 inches from the X-axis. The starting point for this y-scale is 1, and this will be where a transversal cuts y in passing from 1 on x to 1 on z. From this starting point construct the scale whose modulus is $m_y = m_x m_z/(m_x + m_z)$, which is $m_y = (10.73 \times 7.5)/(10.73 + 7.5) = 4.414$ inches. This completes the nomogram for $fd = f \times d$.

5. The construction of $fd^2 = f \times d^2$ is similar. Its form for an addition nomogram is $\log (fd^2) = 2 \log d + \log f$, hence $x = m_1 (2 \log d)$, $z = m_2(\log f)$, $y = m_3[\log (fd^2)]$. If we keep the same outside

axes and scales whose moduli were 10.73 and 7.5 inches for the x- and z-scales respectively, then $m_1 = 10.73/2$ and $m_2 = 7.5$. Substituting in the formula $m_1 m_2/(m_1 - m_2)$ we have $m_3 = 3.14$. Furthermore,

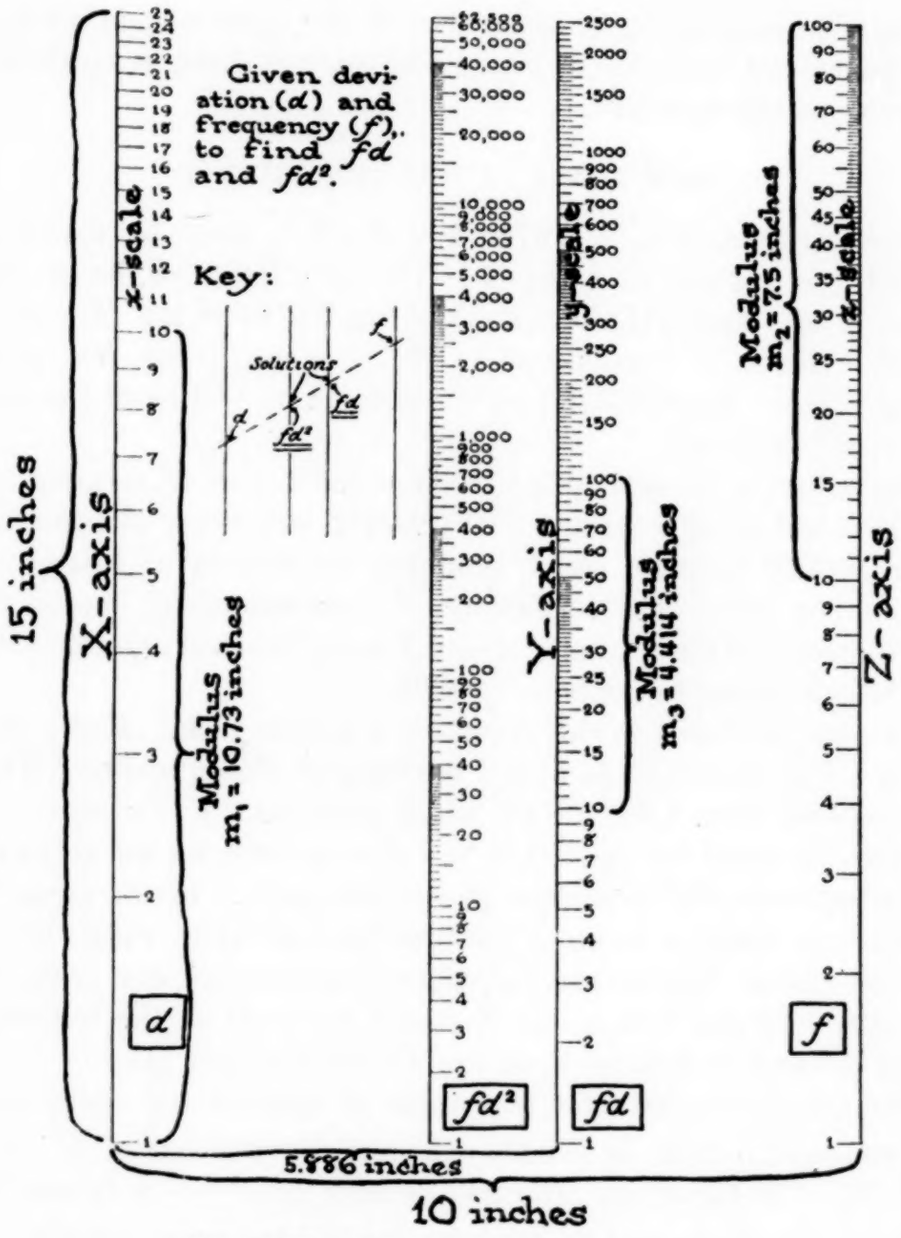

CHART II.—Constructing a nomogram.

the distance of the central axis from the x-axis is $10.73/2/7.5 = x/(10 - x)$, or 4.17 inches.

Adjusting a Nomogram for a Constant.—Where a constant enters into the formula, as in Chart III for computing the probable error of

a correlation coefficient, $PE_r = 0.6745(1 - r^2)/N^{\frac{1}{2}}$, the logarithm of the constant may be added to the sum of the two variable terms by slipping one of the scales up along itself until its point of origin stands above the horizontal level of the point of origin on the other scale by a length corresponding to the value of the constant; or both scales may be moved until the net result of such shifting is equivalent to the value of the constant.

HOW CHART III WAS CONSTRUCTED

The formula, $PE_r = 0.6745(1 - r^2)/N^{\frac{1}{2}}$, may be expressed as $\log (PE_r) = \log 0.6745 + \log (1 - r^2) - \frac{1}{2} \log N$ or, as in the construction of Chart III, $\log (PE_r) - \log 0.6745 = \log (1 - r^2) - \frac{1}{2} \log N$. Hence, $x = m_1[\log (1 - r^2)]$, $z = m_2(\frac{1}{2} \log N)$, and $y = m_3[\log (PE_r) - \log 0.6745]$; and the nomogram will be of the subtraction (division) type.

By referring to any table of values for $1 - r^2$ in terms of r,* we find that values of r from 0.00 to 0.9487 will cover the modulus of $1 - r^2 = 1.00$ to 0.10, which modulus we will fix at fifteen inches. As it is not likely that coefficients of correlation will exceed 0.9487, we will accept it as the limit of our X-axis, but will extend it slightly past fifteen inches to include $r = 0.95$.

Getting the scale on this X-axis is a double task. First, we must obtain a logarithmic scale with a modulus of fifteen inches. This will be numbered from 1.00 to 0.10, which gives us a $1 - r^2$ axis. As we will wish to avoid the operation $1 - r^2$ each time we use this axis, we will incorporate this procedure in our axis scale. So we come to our second step, which is to mark r-values for equivalent values of $1 - r^2$. This operation inverts the apparent direction of our scale, which originally had the 0.10 at the top and the 1.00 at the bottom, and results in the 0 to 0.95 scale on the X-axis of Chart III.

For the Z-axis, we refer to a table of squares and roots and find that values of N from 20 to 2000 will span the modulus of $N^{\frac{1}{2}} = 4.472$ to 44.721. As the X-axis and its modulus for $1 - r^2$ is fifteen inches, we make the Z-axis and its modulus for $N^{\frac{1}{2}}$ the same length. However, as we wish to record values of N, not $N^{\frac{1}{2}}$, on this Z-axis, we will

---

* Miner, J. R.: "Tables of $\sqrt{1 - r^2}$ and $1 - r^2$ for use in Partial Correlation and Trigonometry." Johns Hopkins Press, Baltimore, 1922; Holzinger, K. J.: "Statistical Tables for Students in Education and Psychology." University of Chicago Press, Chicago, 1925.

construct two 7½ inch logarithmic moduli spanning 20 to 2000 on this axis, thus eliminating the need for obtaining the square root of N before using this chart.

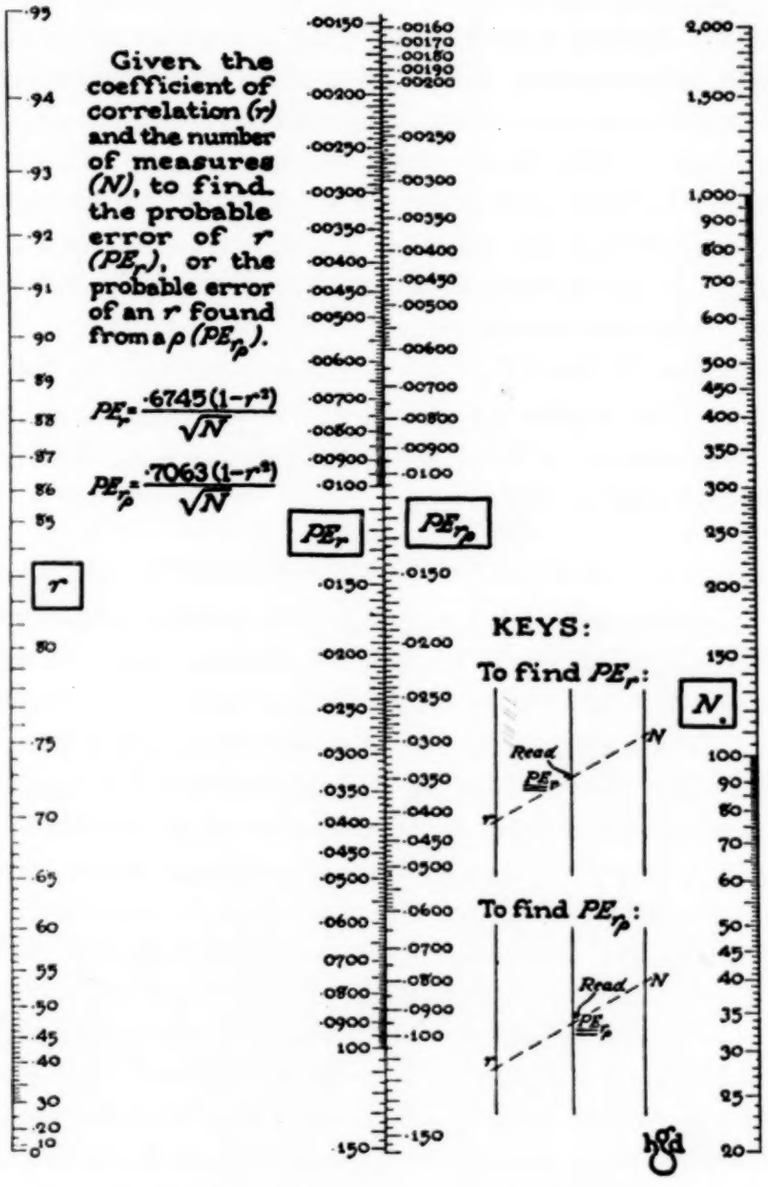
Given the coefficient of correlation (r) and the number of measures (N), to find the probable error of r (PE_r), or the probable error of an r found from a ρ (PE_{r_ρ}).
$PE_r = \frac{.6745(1-r^2)}{\sqrt{N}}$
$PE_{r_\rho} = \frac{.7063(1-r^2)}{\sqrt{N}}$
r
PE_r
PE_{r_ρ}
N
KEYS:
To find PE_r:
Read PE_r
To find PE_{r_ρ}:
Read PE_{r_ρ}

CHART III.—A nomogram for P E$_r$

Thus, as both X- and Z-axes are constructed on fifteen inch moduli, the Y-axis will be erected from a point equidistant from the other two.

The scale on the Y-axis will have a modulus of $m_y = m_x m_z/(m_x + m_y)$, that is $m_y = (15 \times 15)/(15 + 15) = 7.5$ inches. The lower limit (disregarding the constant, 0.6745) would be determined by $r = 0$, therefore, $1 - r^2 = 1$; $N = 20$, therefore, $N^{1/2} = 4.472$;

$(1 - r^2)/N^{1/2} = 1/4.472 = 0.2235$. However, with the constant, 0.6745, included, the lower end of the PE_r-scale on this Y-axis will begin with 0.6745ths of 0.2235, or 0.1508. This point will, of course, be on the Y-axis at a point intersected by an isopleth connecting $r = 0$ and $N = 20$. The scale with modulus of 7.5 inches can then be rapidly constructed, remembering that this is a division (subtraction) nomogram and that the values of PE_r diminish as one goes up the scale.

As a matter of fact, the center-axis on a simple nomogram of this nature can be located and scale points determined by merely noting intersections for the same values of PE_r from two or more values of r and N, using the mathematics given above merely to verify the accuracy of this empirical procedure.

The charting of the PE_{r_ρ}-scale on the other side of the Y-axis can be verified by the reader for practice. Note that its scale is simply the same modulus as in PE_r, allowing for the logarithmic difference between 0.6745 and 0.7063.

FACILITATING CONSTRUCTION

Actual construction of nomograms may be facilitated by the use of semi-logarithmic coordinate papers of various sizes. One needs to be sure, however, that the papers selected have been very accurately made. Sizes of scales on these papers vary from those with a modulus slightly less than two inches up to ten inches. To transfer a scale with a desired modulus from the coordinate paper to the axis, a paper straight-edge is placed on the correct logarithmic paper, and the scale is then copied on this straight-edge and then transferred to the axis. Any size logarithmic scale from one to nearly eighteen inches can be transferred by proper manipulation of the coordinate papers. For example, to obtain a modulus of 12.75 inches, the paper straight-edge is first marked with two points 12.75 inches apart. The straight-edge is then laid diagonally on a ten-inch modulus semi-logarithmic paper with the lower point at 1 and the upper point at 10 on the logarithmic scale. When intermediate markings are properly recorded, the straight-edge contains a logarithmic scale of 12.75 inch modulus. A modulus less than the printed scale is made by a similar process, except that a wooden or metal straight-edge and a triangle are employed to foreshorten a given scale while transferring it from the coordinate paper to the paper straight-edge.

Sometimes a needed modulus will be too long to be copied from the logarithmic papers. Suppose, for example, that in constructing

a certain nomogram that the numbers on a fifteen-inch axis are to be from 100 to 300. The logarithmic difference for these numbers is 0.477. The required modulus, therefore, will be 15/0.477, or 31.45 inches. It will be simpler in this case to make the modulus thirty inches and let the axis be slightly under fifteen inches. Divide thirty inches into ten sections. This will make each section three inches long. On a strip of drawing paper over fifteen inches long and four or more inches wide, draw a center line with an accurate straight-edge. The needed limits of this center line will, of course, be fifteen inches apart. On one side of this center line, beginning with the lower limit usually, lay off five three-inch sections. Number the section lines from 0 to 5. Divide each section into tens, and each of these tens into fives, which is the smallest division practicable for this length modulus.* This side of the center line we will call the "logarithms of the numbers." On the other side of the center line record the numbers corresponding to these logarithms of the numbers, using four-place logarithm tables. Thus, 100 will be opposite 0, 105 will be opposite 0.21, 110 opposite 0.41, 115 opposite 0.61, 120 opposite 0.79, 125 opposite 0.97, 130 opposite 1.14, etc. The writer finds Lacroix and Ragot's tables† to be more convenient for constructing such necessary scales than are the conventional logarithmic tables. The completed logarithmic scale is separated from the "logarithms of the numbers" by cutting down the center line, after which it is transferred to the desired axis.

FOR COMPLEX FORMULAS

The procedure given in this article will enable the reader to construct nomograms for solving simple equations involving two variables and an unknown. For the nomographic solution of more complex formulas, and those with three or more variables, the reader should consult one of the good manuals mentioned in the bibliography.

---

* The writer finds the engineer's triangular ruler which divides the inch into tenths, twentieths, thirtieths, fortieths, fiftieths, and sixtieths, to be exceedingly convenient for decimal division of inches.

† Lacroix, A. and C. L. Ragot: "A Graphic Table Combining Logarithms and Antilogarithms," first edition. The Macmillan Co., New York, 1925.

Contains both four- and five-place graphic tables. The principle is not new. It was used by Gunter before 1620, and was incorporated in Tichy's "Graphische Logarithmentafel" in Austria, and in Farmer's "Logarithm Tables" in England. The arrangement is convenient and reliable. There are no differences to add or subtract, since for every value in one set of units the corresponding value is read directly.

ANNOTATED BIBLIOGRAPHY ON NOMOGRAM CONSTRUCTION

1. "Engineering Index." American Society of Mechanical Engineers, New York.

This, with "Industrial Arts Index" (see below), is the best source for current references on nomograms. Look under the caption "Alignment Charts." The Engineering Societies Library, 29 West 39th St., New York, will supply 11 × 14 inch photostat negatives (paper prints, white printing on black background) of any articles listed in the "Engineering Index" at $.25 for each page plus postage. The Library also makes translations from French and German at $6 per thousand words; with other translating prices obtainable on application.

2. "Cumulative Book Index." H. W. Wilson Company, New York.

This is the "United States Catalog Supplement." Consult under the heading "Graphic Methods," and also "Alignment Charts" and "Nomography."

3. "Education Index." H. W. Wilson Company, New York.

No classification has yet been made for nomograms. Altho the heading "Graphic Methods" is provided, nomograms are not as yet to be found under it; therefore, look under the captions, "Cureton, E. E.," "Dunlap, J. W.," "Griffin, H. D.," and other known designers of educational and psychological nomographs.

4. "Industrial Arts Index." H. W. Wilson Company, New York.

With "Engineering Index," the best source for current material on nomograms. Consult under the captions, "Charts (calculating)," and "Nomography"; also "Alignment Charts," and "Graphic Methods."

5. "International Index of Periodicals." H. W. Wilson Company, New York.

Consult under the captions, "Charts (calculating)," "Graphic Methods," "Nomography (Graphic Methods)," and "Statistical Methods."

6. "Readers' Guide to Periodical Literature." H. W. Wilson Company, New York.

This is devoid of cross references to nomographic literature; consult the other Wilson publications listed.

7. "United States Catalog." H. W. Wilson Company, New York.

Consult especially under "Graphic Methods"; see also, "Alignment Charts," and "Nomography."

8. Almack, J. C. and W. G. Carr: The Principle of the Nomograph in Education. *Journal of Educational Research*, Vol. XIV, 1926, pp. 340–355.

The earliest article I have found in educational literature referring to nomogram construction. Presents nomograms for calculating the intelligence quotient, $IQ = MA/CA$; for predicting school population, $X = ABC$; and for connecting educational need, ability, and performance, $P^2 = 0.81\,NA$.

9. Bergmann, L.: "Nomographische Tafeln für den Gebrauch in der Radiotechnic," second edition. Julius Springer, Berlin, 1926.

Volume eighteen in the Bibliothek des Radio-Amateurs.

10. Brodetsky, S.: "First Course in Nomography," first edition. G. Bell and Sons, London, 1920. First American edition, Open Court Publishing Co., Chicago, 1920.

With Hezlet, the standard British text.

11. Construction and Use of Nomograms. *Engineering and Contracting*, Vol. LV, 1921, pp. 639–641.

Reprinted from the "London Times Engineering Supplement" (date not given). A good exposition of principles and method.

12. CROCKETT, C. W.: The Alinement Chart. *Automotive Industries*, Vol. XLV, 1921, pp. 614–618.

A technical exposition of the mathematical principles underlying nomographic construction.

13. CUMMINGS, S. R. AND J. LIPKA: "Alignment Charts for the Engineer," first edition. Part I, Air and Steam. John Wiley and Sons, New York, 1924.

Unbound, in heavy envelope. Twenty nomograms.

14. CURETON, E. E. AND J. W. DUNLAP: A Nomograph for Estimating a Reliability Coefficient by the Spearman-Brown Formula and for Computing Its Probable Error. *Journal of Educational Psychology*, Vol. XXI, 1930, pp. 68–69.

Interesting to compare with the same formula charted by the writer, reference No. 23.

15. CURETON, E. E. AND J. W. DUNLAP: A Nomograph for Estimating the Reliability of a Test in One Range of Talent When Its Reliability Is Known in Another Range. *Journal of Educational Psychology*, Vol. XX, 1929, pp. 537–638.

A striking contrast in ease of reading a nomogram over an abac (Cartesian diagram) is illustrated by comparing this nomogram with the same formula charted as an abac by P. J. Rulon in the *Journal of Educational Psychology*, Vol. XXI, 1930, pp. 140–142.

16. DALGLEISH, I. S.: "The Lightning Graphs, Series I, General," first edition. Crosby Lockwood and Son, London, 1926.

Ten nomograms are in this first installment of a collection of nomograms intended to assist in frequently recurring calculations that cannot readily be solved with a slide rule. The charts in this installment enable the unknowns to be solved in

$$a^z = b^z + c^z,\ a^z = d,\ \text{and}\ \sqrt[z]{a} = d.$$

These cover relationships between the sides of a right-angled triangle, and give the values of hyperbolic logarithms and reciprocals of all numbers.

17. DEMING, H. G.: The Graphical Interpolation of Tabulated Data. *American Chemical Society Journal*, Vol. XXXIX, 1917, pp. 2388–2392.

Description of a nomogram for interpolating a two page logarithm table so as to provide five-place accuracy, thus saving eighteen pages of logarithm tables. This logarithm table comes with "Nomon," see item 19 below.

18. DEMING, H. G.: The Nomon—a Calculating Device for Chemists. *American Chemical Society Journal*, Vol. XXXIX, 1917, pp. 2137–2144.

A description of elements in the construction of the "Nomon."

19. DEMING, H. G.: "Nomon—A Calculating Chart." Champaign, The University Press, Illinois, 1918.

This nomographic calculating device is equivalent to a slide rule of more than one hundred inches. The writer finds it of great value. The "Nomon" includes a two-page logarithm table with nomographic adjustment to give five-place accuracy. The "Nomon" is sold by the College Book Store, Lincoln, Nebraska, for $1.50.

20. DUNLAP, J. W. AND A. K. KURTZ: "Handbook of Statistical Nomographs, Tables, and Formulas," First edition. World Book Company, Yonkers, 1932.

This is a book that should be in the hands of every statistical worker. It contains twenty-eight valuable statistical nomograms.

21. FELDMAN, W. M.: "Biomathematics; Being the Principles of Mathematics for Students of Biological Science." Charles Griffin and Co., London, 1923.

A valuable introductory manual; discusses nomograms, pp. 118–126. J. B. Lippincott, Philadelphia, is the American agent for all Charles Griffin's publications.

22. FELDMAN, W. M. AND A. J. V. UMANSKI: The Nomogram as a Means of Calculating the Surface Area of the Living Human Body. *The Lancet (London)*, Vol. CCII (Vol. I for 1922), pp. 273–274.

An application in biometry, solving the very tedious DuBois formula,

$$71.84W^{0.425}H^{0.725} = S,$$

in which W = weight, H = height, and S = surface area. I notice a slight mathematical error made by the authors in constructing this nomogram, but not one that seriously affects the accuracy of the chart.

23. GRIFFIN, H. D. Nomogram for Checking the Reliability of Test Scores. *Journal of Applied Psychology*, Vol. XIV, 1930, pp. 609–611.

This chart solves the formulas,

$$r_x = \frac{2r}{1 + r} \text{ and } \bar{r}_x = \frac{nr}{[1 + (n - 1)r]}$$

Compare with the treatment of the latter formula by Cureton and Dunlap, reference 14.

24. GRIFFIN, H. D. Nomograms for Correcting Simple and Multiple Correlation Coefficients. *American Statistical Association Journal*, Vol. XXV, 1930, pp. 316–319.

Nomograms for solving:

$$\bar{r}_{xy} = \frac{(r_{x_1y_2})^{1/2}(r_{x_2y_1})^{1/2}}{(r_{x_1x_2})^{1/2}(r_{y_1y_2})^{1/2}},$$

$$\bar{r}_{xy} = \frac{r_{xy}}{(r_{x_1x_2})^{1/2}(r_{y_1y_2})^{1/2}}, \text{ and}$$

$$\bar{R}^2 = 1 - \frac{1 - R^2}{1 - \frac{n}{N}}.$$

In the third formula, n = number of variables, and N = number of cases.

25. HEDGES, G. L.: Alinement Charts; Explanation of Principles Governing Construction, and Illustrations of Use. *Machinery*, Vol. XXIV, 1918, pp. 615–618.

Gives a systematic method for laying out nomograms by means of an engineer's triangular scale.

26. HEWES, L. I. AND H. L. SEWARD: "Design of Diagrams for Engineering Formulas, and the Theory of Nomography," first edition. McGraw-Hill Book Co., New York, 1923.

Fifty-four illustrative examples are given. Follows Soreau (references 65, 66) in the methods used in charting a formula for nomographic presentation.

27. HEZLET, R. N.: Nomography. "*Encyclopedia Britannica,*" twelfth edition, Vol. XXXI, pp. 1139–1144.

This article was in the new volumes added to the 11th edition in 1922, constituting the twelfth edition; but there is no article on the subject in either the eleventh or thirteenth edition. For the article in the fourteenth edition, see reference 49.

28. How to make a "*Z*" diagram. *Power*, Vol. LXII, 1925, p. 915.

An exposition of a type of diagram that is especially useful when the range of values for the variables is quite limited, and the lowest value is well above zero. See also reference 41.

29. HULL, C. L. A Device for Determining Coefficients of Partial Correlation. *Psychological Review*, Vol. XXVIII, 1921, pp. 377–383.

This was a large alinement board suggested by Kelley's "Tables to Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations." It worked well, but was expensive to manufacture. For more accessible devices see references 31 and 32 or, better still, 71 and 72.

30. KEARTON, W. J. AND G. WOOD: "Alignment Charts for Engineers and Students," first edition. Charles Griffin, London, 1924. J. B. Lippincott, Philadelphia, 1924.

A text and reference book; directions for practical construction; a series of diagrams showing arrangements of lines for equations of three to six variables; thirty-five formulas are charted.

31. KELLEY, T. L. "Chart to Facilitate the Calculation of Partial Coefficients of Correlation and Regression Equations," first edition. School of Education, Special Monograph, No. I. Stanford University Publications, Palo Alto, 1921.

One of the first, if not the first, educational nomograms.

32. KELLEY, T. L. "Alignment Chart of Correlation Functions." Stanford University, Palo Alto, 1921.

Supplement to the preceding monograph, reference 31. A seventeen by twenty-three inch nomogram for solving $1 - r^2$ and $(1 - r^2)^{\frac{1}{2}}$.

33. KONORSKI, B. N. "Die Grundlagen der Nomographie," first edition. Julius Springer, Berlin, 1923.

A technical eighty-nine-page monograph, with interesting applications in many fields.

34. KRAUSE, F.: "Die Nomographie oder Fluchtlinienkunst," first edition. Julius Springer, Berlin, 1922.

A technical fifty-five-page exposition with applications to engineering; author is a Viennese engineer.

35. LACMANN, O.: "Die Herstellung gezeichneter Rechentafeln," first edition. Julius Springer, Berlin, 1923.

Textbook; author is an engineer in Oslo; applications are largely to hydraulics.

36. Les Principes Fondamentaux de la Nomographie. *Le Génie Civil*, Vol. XCV, 1929, pp. 183–184.

A brief exposition.

37. LIESCHE, O.: Nomographie. *Chemische Fabrik*, 1928, pp. 161–164, 228–230, 241–243, 264–265, 280–281, 296–297, 314–315, 359–361, 377, 392–394, 450–451, 463–465, 478, 490–492, 531, 560–561, 583–584, 595–597, 621–623.

For the solution of chemical equations.

38. LIPKA, J.: "Graphical and Mechanical Computation," first edition. John Wiley and Son, New York, 1918.

May be had in one volume or in two parts; Part I is on nomograms, mainly. See also reference 13.

39. LIVINGSTON, R. T.: Practical Construction of Alignment Charts. *Power Plant Engineering*, Vol. XXVII, 1923, pp. 824–826, 873–875, 933–935.

One of the clearest and most practical expositions of empirical construction.

40. LUCKEY, P.: "Einführung in die Nomographie," first edition. B. G. Teubner, Leipzig, 1918, 1920.

A very comprehensive treatise; applications to many fields are suggested.

41. MACKEY, C. O. Graphical Representation of Equations by the *Z*-type of Alignment Chart. *Sibley Journal of Engineering*, Vol. XLIII, 1929, pp. 132–136.

See comment on reference 28.

42. MARSHALL, W. C. "Graphical Methods for Schools, Colleges, Statisticians, Engineers and Executives," first edition. McGraw-Hill Book Co., New York, 1921.

Chapter 8 is on nomography. The book contains a very well analyzed bibliography, pp. 221–248, which, unfortunately, does not give the pages in the volumes for magazine articles, nor the publishers of the books and pamphlets listed.

43. MEHMKE, R.: Numerisches Rechnen. *Encyklopädie der mathematischen Wissenschaften*, Bd. I, Teil II, pp. 938–1079. B. G. Teubner, Leipzig, 1898–1904.

A scholarly and comprehensive discussion of all types of numerical aids and devices. The discussion of calculating machines should be supplemented by Locke, L. L. The History of Modern Calculating Machines, an American Contribution. *American Mathematical Monthly*, Vol. XXXI, 1924, pp. 422–429, in order to supply the contribution of Frank S. Baldwin. Also read Johns, V. On the Mechanical Handling of Statistics. *American Mathematical Monthly*, Vol. XXXI, 1926, pp. 494–502. Mehmke's account of logarithms and the slide rule needs supplementing by Cajori's histories, and articles in *Nature*, Vol. LXXXII, 1909, pp. 267–268, 307–308; Vol. XCIII, 1914, p. 8; and *School Science and Mathematics*, Vol. XX, 1920, pp. 527–530. The French edition of this encyclopedia, "Edition française de l'Encyclopedie des sciences mathematiques," Tome I, has the same article, but with an editing by Maurice d'Ocagne, and, naturally, a more complete exposition of nomography.

44. MEHMKE, R.: "Leitfaden zum graphischen Rechnen," first edition. B. G. Teubner, Leipzig, 1917.

Deals largely with nomograms. Explanations are clear.

45. Nomografía. *Enciclopedia universal ilustrada*, Vol. XXXVIII, pp. 1024–1042. Hijos de J. Espasa, Barcelona, (issued during the 1920's).

An excellent article in this monumental encyclopedia, simply explained and well illustrated. Excellent bibliography, with Spanish, French, German, English, Russian, Italian, and Dutch titles.

46. D'OCAGNE, M.: "*Exposé synthétique des principes foundamentaux de la nomographie,*" first edition. Gauthier-Villars, Paris, 1903.

A short technical and mathematical exposition of the science.

47. D'OCAGNE, M.: "*Calcul graphique et nomographie,*" third edition. Octave Doin, Paris, 1923.

The first edition appeared in 1907–1908.

48. D'OCAGNE, M.: "*Le calcul simplifié par les procédés mecaniques et graphiques,*" second edition. Gauthier-Villars, Paris, 1905.

History and description of calculating machines, slide-rules, tables, abacs, and nomograms. Its discussion of calculating machines and the slide-rule should be supplemented by the references given under 43.

49. D'OCAGNE, M.: Nomography. *Encyclopedia Britannica*, fourteenth edition, Vol. XVI, pp. 483–448.

See also reference 27. The bibliography on p. 484 was evidently compiled in handwriting, and printed without the author's proofreading, for it is full of glaring errors and mis-spellings. The names of Hezlet, Hewes, Lacmann, Soreau, and even of Ocagne himself have been bedeviled.

50. D'OCAGNE, M.: "*Principles usuels de nomographie,*" first edition. Gauthier-Villars, Paris, 1920.

Applications to artillery and aviation.

51. D'OCAGNE, M.: Résumé synthétique des principes fondamentaux de la nomographie. *Revue Génerale Sciences Pures et Appliquées*, Vol. XL, 1929, pp. 325–329.

Synthesis of the methods developed during some forty years.

52. D'OCAGNE, M.: "*Traité de nomographie,*" second edition. Gauthier-Villars, Paris, 1921.

The theory of nomograms with practical applications. The first edition of 1899 has been entirely revised and considerably augmented in this second edition.

53. D'OCAGNE, M.: Esquisse d'ensemble de la nomographie. *Memorial des Sciences Mathematiques*, Vol. IV, 1925, pp. 1–68.

Good bibliography on pp. 64–66.

54. PEARL, R.: "Introduction to Medical Biometry and Statistics," first edition. W. B. Saunders Co., Philadelphia, 1923.

A useful book. Has a section on mathematical formulas and constants, pp. 353–361, providing a handy compendium for use. Discusses nomograms.

55. PEDDLE, J. B.: "The Construction of Graphical Charts," second edition. McGraw-Hill Book Co., New York, 1919.

This was, perhaps, the first American book (first edition, 1910) with simple instructions on nomography. It is a revision and reprint of the author's earlier articles in the *American Machinist*. It is somewhat easier than Lipka, reference 38.

56. VON PIRANI, M.: "Graphische Darstellung in Wissenschaft und Technik," first edition. G. J. Göschen, Leipzig, 1914. Reprint, 1922.

No. 728 in the excellent *Sammlung Göschen*. Deals especially with problems in physics, mechanics, and electricity.

57. ROSE, W. N.: "Mathematics for Engineers." Two vols., first American edition. E. P. Dutton, New York, 1918.

One of the Directly-useful (D.U.) technical series. Has much useful condensed information, such as fixing the decimal point, Vol. I, pp. 6–10, 18–21; nomograms, Vol. I, pp. 429–446, etc.

58. ROSE, W. N.: "Line Charts for Engineers," first American edition. E. P. Dutton, New York, 1923.

Another in the D.U. series. This exposition of nomograms seems to be easier than Lipka's, reference 38, but more difficult than Peddle's, reference 55.

59. RUNGE, C. D. T.: "Graphical Methods," first edition. Columbia University Press, New York, 1912.

A course of lectures delivered at Columbia while Dr. Runge was Kaiser Wilhelm exchange professor. The book is No. 4 of the Ernest K. Adams' Fund for Physical Research, and is frequently bound and cataloged as such. Theory is emphasized.

60. RUNGE, C. D. T.: "Graphische Methoden," second edition. B. G. Teubner, Leipzig, 1919.

This is an augmenting of the earlier book in English of the same name, reference 59. The first edition appeared in 1913. Chapter 2 has an exposition of nomograms. Like reference 59, the treatment of the subject matter is technical.

61. SCHENKEL, H.: Die Nomographie in der E T Z bis 1924. *Elektrotechnische Zeitschrift* (*E T Z*), Vol. XLVI, 1925, pp. 1060–1061.

Bibliography of thirty-seven items from Vol. XXVIII, 1907 through Vol. XLV, 1924. See also reference 69.

62. SCHILLING, F.: "Über die Nomographie von M. d'Ocagne," second edition. B. G. Teubner, Leipzig, 1917.

The first edition, of which this is a revision, appeared in 1900.

63. SCHWERDT, H.: "Lehrbuch der Nomographie," first edition. Julius Springer, Berlin, 1924.

One of the best textbooks for serious students of the science; theory based on geometry. Valuable suggestions for assuring accuracy in charting.

64. SEWARD, H. L.: Graphical Methods of Calculation. *Mechanical Engineering*, Vol. XLVII, 1925, pp. 1019–1033.

"The construction of Cartesian diagrams and of alignment charts for use in the rapid determination of unknown quantities in mathematical expressions." This paper was presented and discussed at the 1925 meeting of the American Society of Mechanical Engineers. A summary of the discussions is found in *Mechanical Engineering*, Vol. XLVIII, 1926, pp. 592. I consider this the best exposition in English of the methods of construction advocated by Soreau (references 65 and 66).

65. SOREAU, R.: Contribution à la théorie et aux applications de la nomographie. *Extrait des Mémoires et Compte Rendu des Travaux de la Société des Ingénieurs Civils de France*, August bulletin, No. 8, 1901, 2d vol., pp. 191–512.

Soreau uses determinants in organizing a formula for presentation by nomograms; see also references 26, 64, and 66.

66. SOREAU, R.: *Nomographie, ou traite des abaques*, first edition. Two vols, Etienne Chiron, Paris; 1921.

A recent exposition of Soreau's principles.

67. STRACHAN, R. C.: Nomographic solutions for formulas of various types. *Proceedings of the American Society of Civil Engineers*, Vol. XL, 1914, pp. 2975–2996; *Transactions of the American Society of Civil Engineers*, Vol. LXXVIII, 1915, pp. 1359–1408.

A good presentation of nomograms with a discussion by five other engineers. Abstracts from these articles are found in *Engineering Record*, Vol. LXXI, 1915, pp. 807–809.

68. SWETT, G. W.: "Construction of Alignment Charts," first edition. John Wiley and Sons, New York, 1928.

An excellent little manual.

69. WEHAGE, D.: Neuere Anwendungen der Nomographie (Mitteilung aus dem Telegraphentechnischen Reichsamt). *Elektrotechnische Zeitschrift* (*E T Z*) Vol. XLV, 1924, pp. 1342–1346.

A good bibliography of current articles on nomography; mostly in German, but some from other languages. See also reference 61.

70. WERKMEISTER, P.: "Das Entwerfen von graphischen Rechentafeln (Nomographie)," first edition. Julius Springer, Berlin, 1923.

A good manual on nomography, with a good bibliography. The author has been a contributor on the subject to many German mathematical and technical journals.

71. WOOD, E. R.: "A Graphic Method of Obtaining the Partial-correlation Coefficients and the Partial-regression Coefficients of Three or More Variables," first edition. Supplementary Educational Monograph, No. 37. The University of Chicago, Chicago, 1931.

This very excellent chart for solving coefficients of partial correlation and of partial regression by a combined abac and nomographic method was completed in 1923 but was not published until recently. It is especially valuable in its treatment of the coefficients of partial correlation, and anyone having many of these to solve would do well to obtain this monograph and the forms mentioned in reference 72.

72. WOOD, E. R.: "Forms to Facilitate the Calculation of Partial-correlation Coefficients, Partial-regression Coefficients, and Partial-regression Equations." Published by E. R. Wood, State Department of Education, Columbus, Ohio.

The best abac-nomographic devices yet constructed for these purposes.

73. WOOD, E. R.: "Chart for Obtaining Correlation Coefficients." Published by E. R. Wood, State Department of Education, Columbus, Ohio.

An abac-nomographic device to solve a formula of the type

$$r = \frac{w}{v^{\frac{1}{2}} z^{\frac{1}{2}}}$$

INTERPRETATION OF CORRELATION ON THE BASIS OF COMMON ELEMENTS

P. H. NYGAARD

North Central High School, Spokane, Washington

The whole theory of correlation as used in statistics has been built around the regression equation, the coefficient of correlation being derived from the regression coefficient. There is, however, another interpretation of correlation which tends to make clearer just what correlation implies. Its application on a wider scale than heretofore would seem to be desirable. According to this interpretation any trait can be regarded as composed of a summation of smaller subtraits, and these again may consist of still smaller components. The final components into which a trait is divided are called the elements composing that trait. The coefficient of correlation between any two traits will then depend upon the number of elements common to the two traits.

This explanation of correlation in terms of common elements was first proposed by Weldon. It is dealt with fully by Brown and Thomson.[1] A summary of it is given by Kelley.[2] The part of this paper devoted to the statement of the theory contains nothing basically new, although the writer has made some modifications both in the assumptions and in the notation. The parts devoted to the illustration, the proof of agreement with the regression method, and the application to the Spearman two factor theory are original contributions as far as the writer knows.

I. STATEMENT OF THE THEORY

Any trait is made up of a number of small elements. It is assumed that all of these elements are equally important and equally probable. By equally important is meant that the presence of any one element has just as much effect upon the measure of a trait as does the presence of any other element, or, in other words, that any measure of a trait depends directly upon the number of elements of that trait present. Thus, if arithmetic ability is the trait considered, a person's measurement in arithmetic ability will depend upon the number of elements of

[1] Brown, William and Godfrey H. Thomson: "Essentials of Mental Measurement." Cambridge, 1921, p. 176.

[2] Kelley, Truman L.: "Statistical Method." The MacMillan Co., 1923, p. 189.

arithmetic ability that he possesses. By equally probable is meant that every element of a trait is just as likely to be present as absent.

These elements do not occur in variable intensity. An element is either present or absent, never just partly present. If the traits in question are mental, the elements are comparable to, or may be conceived to be dependent upon, the association bonds of psychology. In dealing with such traits the presence or absence of an element may be interpreted to mean that a certain association bond has been formed or has not been formed, or, more broadly, that the specific process represented by the element has been developed or has not been developed. Psychologists claim that the transfer of training from one mental activity to another occurs only to the extent that there are common elements involved. These common elements are precisely the same as those postulated in this paper. The theory is, however, adapted equally well to traits that are not mental.

Let A and B be any two traits. Represent the number of elements composing trait A by n_a and the number of elements composing trait B by n_b. If a large number of random measures of a trait are made, it follows from the equi-importance and the equi-probability of the elements that the arithmetic mean of the measures will contain $\frac{1}{2}$ of the total number of elements that make up the trait. That is,

$$\begin{aligned} M_a &= \tfrac{1}{2}n_a \\ M_b &= \tfrac{1}{2}n_b \end{aligned} \qquad (1)$$

It can also be shown that the standard deviation of the measures of any trait will be $\frac{1}{2}$ of the square root of the number of elements composing the trait.[1] That is,

$$\begin{aligned} \sigma_a &= \tfrac{1}{2}\sqrt{n_a} \\ \sigma_b &= \tfrac{1}{2}\sqrt{n_b} \end{aligned} \qquad (2)$$

If there is no correlation between the two traits, every element of trait A will be different from every element of trait B. In case correlation exists, a number of the elements of A are also elements of B. The number of such common elements will be represented by n_{ab}. It may then be proved that the product-moment coefficient of correlation is equivalent to the ratio of the number of elements that the two traits have in common to the geometric mean of the number of elements in each trait. That is,

[1] Kelley: "Statistical Method." P. 190. The fact that the measures there used are deviations from their respective means will not affect their standard deviations.

$$r_{ab} = \frac{n_{ab}}{\sqrt{n_a \cdot n_b}} \tag{3}$$

For the special case in which both traits are composed of the same number of elements, that is $n_a = n_b$, it is obvious that the coefficient of correlation is simply the ratio of the number of common elements to the total number of elements in each trait. That is,

$$r_{ab} = \frac{n_{ab}}{n_a} = \frac{n_{ab}}{n_b}, \tag{4}$$

when $n_a = n_b$. Thus, if it is assumed that the yield per acre of wheat depends upon the same number of elements as the yield per acre of barley, a correlation coefficient of .80 between the average wheat yield and the average barley yield in a certain community over a term of years would be interpreted to mean that 80% of the elements that determine the wheat yield are also elements that determine the barley yield.

II. ILLUSTRATION

Let us assume that trait A is made up of two elements, A_1 and A_2, and that trait B is composed of three elements, B_1, B_2, and B_3. Any trait would theoretically consist of a large number of elements, but we shall assume that the elements are few so as to make the illustration simpler. Let A_1 and B_1 be identical. We then have $n_a = 2$, $n_b = 3$, and $n_{ab} = 1$. Formula (3) gives $r_{ab} = \frac{1}{\sqrt{2 \cdot 3}} = \frac{1}{\sqrt{6}}$. We shall show that this agrees with the product-moment coefficient obtained directly from the probable distribution of the measures of A and B as presented in the accompanying table.

Assume that we are dealing with 16 individual measures of trait A. Since each of the two A elements is just as likely to be absent as present, out of the 16 measures of A the probability is that 4 will have no A element present, which is indicated in the table by A_0. It is probable that 8 of the 16 will contain one A element, 4 of which will be A_1 and the other 4 will be A_2. The remaining 4 measures will then contain both A_1 and A_2. Considering next the 4 measures that have A_0, it is probable that one of the corresponding measures of B will have no B element present, one will have the B element B_2, one the B element B_3, and one the B elements B_2 and B_3. None of these can have B_1, for A_1 was not present. When A_1 alone is present, the probable arrangement of the corresponding B measures is as

shown. Since A_1 is present, B_1 must be present in all of them. Similarly for the rest of the second column. Notice that both of the distributions are in exact agreement with the laws of probability.

A elements present	B elements present	Measures of A	Measures of B	x	y	xy	x^2	y^2
A_0	B_0	0	0	−1	−1.5	1.5	1	2.25
A_0	B_2	0	1	−1	−.5	.5	1	.25
A_0	B_3	0	1	−1	−.5	.5	1	.25
A_0	B_2B_3	0	2	−1	.5	−.5	1	.25
A_1	B_1	1	1	0	−.5	0	0	.25
A_1	B_1B_2	1	2	0	.5	0	0	.25
A_1	B_1B_3	1	2	0	.5	0	0	.25
A_1	$B_1B_2B_3$	1	3	0	1.5	0	0	2.25
A_2	B_0	1	0	0	−1.5	0	0	2.25
A_2	B_2	1	1	0	−.5	0	0	.25
A_2	B_3	1	1	0	−.5	0	0	.25
A_2	B_2B_3	1	2	0	.5	0	0	.25
A_1A_2	B_1	2	1	1	−.5	−.5	1	.25
A_1A_2	B_1B_2	2	2	1	.5	.5	1	.25
A_1A_2	B_1B_3	2	2	1	.5	.5	1	.25
A_1A_2	$B_1B_2B_3$	2	3	1	1.5	1.5	1	2.25

$$\Sigma xy = 4;\ \Sigma x^2 = 8;\ \Sigma y^2 = 12.$$

$$r_{ab} = \frac{4}{\sqrt{8 \cdot 12}} = \frac{4}{\sqrt{96}} = \frac{4}{4\sqrt{6}} = \frac{1}{\sqrt{6}}.$$

The numbers in the third and fourth columns give the measures of trait A and trait B for each of the 16 individuals. The numbers are obtained directly from the first two columns, for all the elements are assumed to be equally important. The rest of the table is simply the process of calculating the product-moment coefficient of correlation, which is found to be exactly the same as obtained by the use of formula (3).

III. PROOF OF AGREEMENT WITH THE REGRESSION THEORY

The common element interpretation of correlation does not conflict with the usual interpretation based on the regression coefficient. The explanation of correlation on the basis of common elements, as here presented, gives the same results as, and is in some respects simpler than, the usual explanation. In our proof of this we shall obtain a regression equation in terms of the common element theory,

and then show that this is equivalent to the conventional regression equation.

As before, let trait A be composed of n_a elements, trait B of n_b elements, and let the number of elements common to the two traits be represented by n_{ab}. Let X_a represent a particular measure of trait A, which, since the elements that make up trait A are all equally important, amounts to the same as saying that X_a represents the number of A elements present in this particular measure. Some of the elements composing X_a will be elements common to A and B. The number of such elements will be the total number of elements common to A and B multiplied by the ratio of the number of elements in X_a to the total number of A elements. This follows from the equi-probability of the elements. Accordingly, this particular measure of A will be made up of $\frac{X_a}{n_a} \cdot n_{ab}$ common elements plus a number of specific A elements. The corresponding particular measure of trait B will therefore have $\frac{X_a}{n_a} \cdot n_{ab}$ common elements plus a number of specific B elements. The total number of specific B elements is $n_b - n_{ab}$. The probability is that $\frac{1}{2}(n_b - n_{ab})$ of these will be present in this particular measure. Hence we obtain

$$X_b = \frac{X_a}{n_a} \cdot n_{ab} + \tfrac{1}{2}(n_b - n_{ab}).$$

This gives

$$X_b = \frac{X_a \cdot n_{ab}}{n_a} - \tfrac{1}{2}n_{ab} + \tfrac{1}{2}n_b,$$

or

$$X_b = \frac{X_a \cdot n_{ab}}{n_a} - \frac{n_a \cdot n_{ab}}{2n_a} + \tfrac{1}{2}n_b.$$

From this is obtained the very simple regression equation in terms of the common element theory,

$$X_b = \frac{n_{ab}}{n_a}(X_a - \tfrac{1}{2}n_a) + \tfrac{1}{2}n_b. \qquad (5)$$

By making further changes in formula (5) we get

$$X_b = \frac{n_{ab}}{\sqrt{n_a \cdot n_b}} \cdot \frac{\frac{1}{2}\sqrt{n_b}}{\frac{1}{2}\sqrt{n_a}}(X_a - \tfrac{1}{2}n_a) + \tfrac{1}{2}n_b.$$

But by formula (3), $\frac{n_{ab}}{\sqrt{n_a \cdot n_b}} = r_{ab}$; by formula (2), $\frac{1}{2}\sqrt{n_a} = \sigma_a$, and

$\frac{1}{2}\sqrt{n_b} = \sigma_b$; by formula (1), $\frac{1}{2}n_a = M_a$, and $\frac{1}{2}n_b = M_b$. Substituting these values, we obtain

$$X_b = r_{ab} \cdot \frac{\sigma_b}{\sigma_a}(X_a - M_a) + M_b.$$

This is the conventional regression equation used in predicting a measure of B from a given measure of A. This equation has been derived by using the assumptions and formulas of the common element theory. That this can be done is sufficient proof of the inherent soundness of these assumptions and formulas.

IV. APPLICATION TO THE SPEARMAN TWO-FACTOR THEORY

The Spearman two-factor theory can be readily substantiated by means of the common element interpretation of correlation. It will be necessary first to summarize the Spearman procedure. Let A, B, C, and D be four traits. The quantity $r_{ab} \cdot r_{cd} - r_{ac} \cdot r_{bd}$ is called the tetrad difference.[1] Each of the traits is divided for analysis into the sum of a general factor called g plus a specific factor. Thus, X_a, any particular measure of A, is composed of $r_{ag} \cdot g$ plus a specific factor; X_b is similarly composed of $r_{bg} \cdot g$ plus a specific factor; X_c is composed of $r_{cg} \cdot g$ plus a specific factor; and X_d is composed of $r_{dg} \cdot g$ plus a specific factor. The g is called the general factor because it is common to all four traits, although it enters into them in varying amounts. The Spearman technique is based on this theorem:

> If each of the four traits under consideration can be divided as shown above into a general factor plus a specific factor, and if every specific factor is altogether different from g and from every other specific factor, then the tetrad difference equals 0.

This theorem will now be proved by means of the common element interpretation of correlation.

The measure of trait A is composed of a common factor, $r_{ag} \cdot g$, plus a specific factor. If the factor concept is replaced by that of elements, the amount of g will be replaced by the number of g elements, so that X_a will be composed of $r_{ag} \cdot n_g$ elements common to A and g plus a number of specific elements. The number of g elements in X_a depends, it is seen, upon the value of the fraction r_{ag}. Similarly, the number of g elements in X_b depends upon the fraction r_{bg}. Accord-

[1] There are, of course, two other tetrad differences involving these four traits, but what is stated here will apply just as well to the other two.

ing to the hypothesis of the theorem being proved, none of the specific elements can be common to X_a and X_b. Hence the only elements of X_a that are also in X_b are those g elements that are found in both X_a and X_b. The probable number of such elements common to X_a and X_b is the product of the respective fractions times n_g. Therefore, in our notation, $n_{ab} = r_{ag} \cdot r_{bg} \cdot n_g$. By the same reasoning $n_{cd} = r_{cg} \cdot r_{dg} \cdot n_g$, $n_{ac} = r_{ag} \cdot r_{cg} \cdot n_g$, and $n_{bd} = r_{bg} \cdot r_{dg} \cdot n_g$. From these four equations it is obvious that $n_{ab} \cdot n_{cd} = n_{ac} \cdot n_{bd}$. Dividing each side of this equation by $\sqrt{n_a \cdot n_b \cdot n_c \cdot n_d}$ gives

$$\frac{n_{ab}}{\sqrt{n_a \cdot n_b}} \cdot \frac{n_{cd}}{\sqrt{n_c \cdot n_d}} = \frac{n_{ac}}{\sqrt{n_a \cdot n_c}} \cdot \frac{n_{bd}}{\sqrt{n_b \cdot n_d}}.$$

By formula (3) this becomes

$$r_{ab} \cdot r_{cd} = r_{ac} \cdot r_{bd},$$

or

$$r_{ab} \cdot r_{cd} - r_{ac} \cdot r_{bd} = 0.$$

Spearman does not claim that the converse of this theorem is necessarily true.[1] If the tetrad difference is known to equal 0, the four traits can be divided in other ways than the pattern given by the two-factor theory. Spearman does claim, however, that the two-factor pattern is the simplest and is the one most pregnant with possibilities in the field of psychological research. Kelley,[2] on the other hand, prefers to think of the variables as made up of an infinite number of elements even when the tetrad difference equals 0.

The common element interpretation of correlation may throw some light on this converse problem. If the tetrad difference equals 0, we have the tetrad equation,

$$r_{ab} \cdot r_{cd} - r_{ac} \cdot r_{bd} = 0.$$

By formula (3) this becomes

$$\frac{n_{ab}}{\sqrt{n_a \cdot n_b}} \cdot \frac{n_{cd}}{\sqrt{n_c \cdot n_d}} - \frac{n_{ac}}{\sqrt{n_a \cdot n_c}} \cdot \frac{n_{bd}}{\sqrt{n_b \cdot n_d}} = 0.$$

Clearing of fractions gives

$$n_{ab} \cdot n_{cd} - n_{ac} \cdot n_{bd} = 0.$$

---

[1] Spearman, C. E.: *Journal of Educational Psychology*, February, 1931, pp. 113 and 114.

[2] Kelley, T. L.: "Crossroads in the Mind of Man." Stanford Press.

From this is obtained the following equivalent of the tetrad equation in terms of the common element theory:

$$\frac{n_{ab}}{n_{ac}} = \frac{n_{bd}}{n_{cd}}. \tag{6}$$

From this it follows that if the tetrad difference equals 0, the four traits must be constituted in such a way that the number of elements common to B and A is to the the number of elements common to C and A as the number of elements common to B and D is to the number of elements common to C and D. The two-factor pattern meets this criterion, as we have shown, but it is obvious that other patterns based on the summation of elements will also satisfy it.

RELIABILITY OF GROUP DISTINCTIONS

RAYMOND FRANZEN AND MAHEW DERRYBERRY

American Child Health Association, New York City

If we correlate two estimates of a given quality each of which has served to distinguish the same individuals (or groups) we commonly call this a reliability coefficient. This coefficient states the degree to which estimates agree when two different samplings of the same type of materials are used. It is a measure of the precision of the test *materials.*

Reliability may, however, have an entirely different meaning. It may apply to the precision with which groups are differentiated with a given set of materials. Tests which give reliable indications of individual differences may, nevertheless, fail to distinguish between groups when these groups are chosen in some particular manner. Average scores on a test for random halves of groups may fail to correlate because discriminations made by the test are not rooted in the manner in which the groups are selected. Correlation between *average scores for random halves of groups on one set of materials* tests the precision with which distinctions are made between these groups with these materials.

Whether or not distinctions between the units measured are true discriminations depends, first, upon the common factor in the sampling of test materials and, second, upon the common factor in the sampling which controls the composition of the units. When the units are individuals this latter sampling is of different times and conditions. When the units are groups of individuals it is a sampling of individuals belonging to the units as defined.

THE RELIABILITY OF DISTINCTIONS BETWEEN GROUPS

Reliability of the first sort is that which is commonly reported for a test. It deals with the distinctions made between items of the population—not, however, in units of the test but in multiples of the σ of the distribution. To distinguish means reliably we must know that these means are a given distance apart on two separate samplings of material and *this distance must be measured in multiples of the obtained σ of the means* or an estimate thereof not suffering from systematic bias.

To distinguish means that are as far apart as $.26\sigma_\infty$ (the σ_∞ being the true σ in a distribution of individuals) the reliability of the test

required is only .50.[1] But to distinguish means that are as far apart as $.26\sigma_\infty$ (the σ_∞ being the true σ of averages) the reliability of the test required is .94. It takes just as reliable a test to distinguish two averages *in terms of the true σ of the averages* as it does to distinguish two individuals *in terms of the true σ of the individuals.* But of course the reliability involved in the distinction of means is the correlation of means computed from random halves of the material. This correlation may be higher than that between random halves of the material among individuals *but it is not necessarily so.*

It is agreed that group averages may be distinguished a given distance (measured in σ of individuals) with a test that would be unreliable as a tool for individual discrimination. But many problems demand reliable group distinctions *in terms of the variability between groups.* It is this latter type of problem which interests us here. Sociological problems deal with such variables as "the average twelve-year-old Thorndike-McCall scores" in a hundred schools or "the five-year-old diphtheria mortality" in a hundred cities; and then, of course, we need to know the reliability of the distinctions with which we are dealing. Social investigations are rife with variables like these. There is a false impression that the distinctions between these averages are always more dependable than distinctions made when individuals are the units of the investigation. This leads to the acceptance of unreliable measures and unfortunate conclusions.

As early as 1913 Pearson[2] showed that $r_{\bar{x}\bar{y}} = r_{xy}$ where $\bar{x}$ and $\bar{y}$ are the fluctuations in the means of the several samplings. In a concrete case this means that if five thousand individuals are divided into fifty random groups of one hundred cases each, then the correlation between *averages for these groups computed from random halves of the materials* is the same as the reliability coefficient of the five thousand cases when the same division of materials is used. Of course, this assumes that the population used is the same in both cases or that both populations are representative random samplings.

For instance, the correlation between two forms of a college entrance test using five thousand entering freshmen from fifty colleges would be the same as the correlation of these two forms using averages of the fifty entering groups provided selection of college was unrelated to the

---

[1] As shown in Kelley, T. L.: "Interpretations of Educational Measurements." Pp. 210ff.

[2] Pearson, Karl: On the Probable Error of Frequency Constants. *Biometrika*, Vol. IX, 1913, pp. 1–10.

quality measured by the test. However, correlation between Form 1 and Form A of a test is less for individuals than for means if the groups are chosen in some manner related to the quality being measured by the test. If, for instance, the fifty colleges had different standards for entrance such that the one hundred entrants that made the highest scores were selected by one college and the next one hundred were selected by a second college, etc., then the correlation between means of Form 1 and means of Form A would be considerably higher than is the case between these two forms in the universe of individuals. Therefore the degree to which distinction between groups with these materials is more reliable than distinction between individuals is dependent upon the degree to which general ability of entrants is related to the choice of college. But this is the very problem with which we began—namely, the distinction between groups.

It is clear that we cannot assume that the reliability is the same or any defined amount greater than is the case with individuals.[1] We must determine what it is. If the correlation between averages of Form 1 with averages of Form A is low then distinctions between groups are unreliable. The false logic involved in the belief that an unreliable test may, nevertheless, be used to measure a group is rooted in the illusion of test units.

The errors of estimate made by a test with individuals is given by

$$\sigma_{x_1 \cdot \infty} = \sigma_{x_1}\sqrt{1 - r_{x_1 x_2}} \qquad (1)^2$$

where x_1 and x_2 represent the scores on two comparable tests. The errors of estimate made by the same test with groups is given by

$$\sigma_{\bar{x}_1 \cdot \infty} = \sigma_{\bar{x}_1}\sqrt{1 - r_{\bar{x}_1 \bar{x}_2}} \qquad (2)^3$$

but when $r_{\bar{x}_1 \bar{x}_2} = r_{x_1 x_2}$ the above formula becomes

$$\sigma_{\bar{x}_1}\sqrt{1 - r_{x_1 x_2}} \qquad (2a)$$

From this it can be seen that the error of measurement for averages of groups is always less than the error of individual measurement *in terms of scores on the test.* But the scores of the test are not the proper relativity. The reliability of distinctions rendered between groups or individuals is determined by the error of discrimination.

---

[1] Though, of course, we can assume, barring any fantastic manner of selection, that the reliability is at least as large as it is with individuals.

[2] Kelley, Truman Lee: "Interpretation of Educational Measurements." Pp. 171ff.

[3] $\sigma_{\bar{x}} = \sigma_x/\sqrt{n}$.

When the discrimination with which we are concerned is between two particular means then this depends upon the magnitude of the difference in question divided by the standard error of this difference. When, however, we are concerned with the whole range of distinctions as we are when we correlate means on one variable with means on another, then the best statement of reliability is the ratio of errors of measurement to the variability under consideration. For groups this ratio is

$$\frac{\sigma_{\bar{x}_1}\sqrt{1 - r_{\bar{x}_1\bar{x}_2}}}{\sigma_{\bar{x}_1}} = \sqrt{1 - r_{\bar{x}_1\bar{x}_2}}$$

and for individuals it is

$$\frac{\sigma_{x1}\sqrt{1 - r_{x_1x_2}}}{\sigma_{x_1}} = \sqrt{1 - r_{x_1x_2}}$$

Unless we can show a higher reliability between groups we see that the ratio is the same whether we use individuals or groups. It follows then that a test which is not reliable enough to distinguish between individuals is not reliable enough to distinguish between groups unless we control the selection of constituents. Relative to the inch the average height of one hundred ten-year-olds has less error than has the measured height of any one ten-year-old. But relative to other random groups of one hundred ten-year-olds it has as much error as has the measure of one ten-year-old relative to a sampling of his peers. We rarely place the average of a group in a distribution of individuals. We compare it to the averages of other groups,—and the best measure of the reliability of a measure in making distinctions between groups is $\sqrt{1 - r_{11}}$.

THE RELIABILITY OF THE SAMPLING OF INDIVIDUALS COMPOSING GROUP UNITS

Whenever distinctions between groups are sought we are concerned with the criteria of grouping. In addition to knowing that the sampling of materials composing the tests are adequate we need to know that the sampling of individuals composing the groups are truly representative of their designations. This becomes essential where some of the variables treated are based on these same designations and not upon the sampling of individuals composing the basis of the other measures. If, for instance, we wish to correlate the size of colleges with the average score of entrants or some test of the real estate values of communities with the intelligence of twelve year old children in the

public schools, then the value of the results hinges upon the representative nature of the groups. Discriminations must be due to the criteria of grouping and not to chance inclusion of cases since other distinctions are directly dependent upon these criteria.

In the School Health Study[1] the data were sixty-eight fifth and sixth grade groups. In some measures the distinctions, using one chance half of each group, are practically identical with the distinctions using the other half. This is true of a measure of "Cleanliness" based upon inspection for ingrained dirt. The correlation representing homogeneity in groups is .94. The same coefficient for percentage of vision defects corrected is less than .30. In this latter case the differences between groups are not rooted in the manner in which the groups were chosen. Since the units as chosen designate a community we may conclude that the differences between groups are not due to the community and therefore may not be attributed in any way to differences which exist between the communities as chosen. No correlation may be expected between this measure and aspects of public health which differ from community to community.

This reliability coefficient which measures the adequacy of the criteria of grouping may be low even if the correlation between two tests made on the same individuals is high. The test of whether vision defects in individuals are corrected or not is very dependable. But school groups do not differ in this quality in a way that is rooted in that kind of grouping. Suppose we measured the general intelligence of all children named Brown, Jones, Williams and forty-seven other Anglo-Saxon surnames and then, using the fifty groups, obtained two scores on each by dividing each one into half by chance selection. We would probably get a very low correlation meaning that the groups were not distinguished in intelligence. Suppose, however, the groups were so chosen in terms of surname that each one represented a different set of immigration circumstances. Then the same kind of correlation would probably be high. This type of reliability is fully as important in group distinctions as is the reliability of the materials.

CATEGORICAL DISTINCTION WITHIN THE GROUPINGS

Groups as chosen may have some categorical aspect like sex or race. The method of calculating the correlation between random

---

[1] A research on school health programs conducted by the Research Division of the American Child Health Association.

halves of the groups must then be modified to include consideration of evidences of heterogeneity between sub-groups. Such a procedure does not aim at differences between sub-groups that are correlated with the differences between total groups but at differences which change the relation of groups as designated when viewed from one or the other category. This may be clearly seen if we take an extreme case.

Let X_a = Average measure of a random half of the boys of a group in a given quality.
X_b = Average measure of a second random half of the boys of a group in the same quality.
X_c = Average measure of a random half of the girls of a group in the same quality.
X_d = Average measure of a second random half of the girls of a group in the same quality.

Let us assume that the influences due to differences between communities in the given quality are of one kind on boy scores and entirely another kind on girl scores. The average boy score measures quite different influences from the average girl score. Let us further assume that in this quality the correlations of boys with boys and of girls with girls are unity. Then by definition the intercorrelations of these four qualities become:

	X_a	X_b	X_c
X_b	1.00		
X_c	.00	.00	
X_d	.00	.00	1.00

In this extreme case it can be shown that the correlation between random halves with the same relative sex representation is unity whereas the correlations of boy scores with girl scores are zero. Naturally a measure like this one must be so treated as to yield for each group a score for each sex since scores based on totals conceal true differences appertaining to one or the other sex.

Of course in our usual data we do not find any such extreme pattern. However, materials treated in this manner allow a common sense decision of degree to which group scores belonging to one category show enough individuality in dispersion to make separate treatment of sub-groups necessary. The following examples will illustrate:

RELIABILITIES OF PROPORTION OF MISSING FIRST MOLARS AMONG SIXTY-EIGHT FIFTH AND SIXTH GRADE GROUPS

	Boys *A*	Boys *B*	Girls *A*
Boys *B*	.67		
Girls *A*	.48	.59	
Girls *B*	.65	.62	.72

RELIABILITIES OF PROPORTION OF CARIES CORRECTED AMONG SIXTY-EIGHT FIFTH AND SIXTH GRADE GROUPS

	Boys *A*	Boys *B*	Girls *A*
Boys *B*	.80		
Girls *A*	.78	.87	
Girls *B*	.77	.83	.87

In the first of these cases it is safer to treat the data by sex groups. Deviations from the average occur sufficiently often for one sex without a parallel deviation in the other to justify separate treatment. In the second case it is entirely safe to treat both sexes together.[1] The difference between these two sets of materials is most readily observable when combinations of groups bring out, first, the correlation of boys with girls and, second, the correlations with the two possible mixed groups.[2]

	Materials showing sex difference	Materials free from sex difference
Boy with girl	.69	.89
Groups with the same relative sex representation	.86	.92
	.78	.91

[1] This, of course, assumes that the materials are corrected for differences in sex average in relation to the relative frequency of the sexes.

[2] These combinations can be made with the formula:

$$r_{(x+y)(z+w)} = \frac{r_{xz}\sigma_x\sigma_z + r_{xw}\sigma_x\sigma_w + r_{yz}\sigma_y\sigma_z + r_{yw}\sigma_y\sigma_w}{\sqrt{\sigma_x^2 + \sigma_y^2 + 2r_{xy}\sigma_x\sigma_y}\sqrt{\sigma_z^2 + \sigma_w^2 + 2r_{zw}\sigma_z\sigma_w}}$$

SUMMARY

1. The precision of the distinctions made with a given set of test materials is not always better for groups than it is for individuals. The reliability of the materials with different kinds of groupings must be determined.

2. The correlation between random halves of constituent groups is an important aspect of the reliability of group distinctions.

3. Categorical divisions in a series of groups must be examined for unique effect upon dispersion.

RATING TENDENCIES OF PERSONALLY SELECTED JUDGES

RICHARD STEPHEN UHRBROCK

The Procter and Gamble Company, Ivorydale, Ohio

I. INTRODUCTION

Judgments concerning personal characteristics have been expressed by word of mouth; in personal letters; in communications addressed "to whom it may concern"; and by means of simple or elaborate checking schemes and rating devices. The problem of obtaining valid judgments of personal characteristics for the purposes of vocational placement is extremely difficult from several points of view. Interviews with judges (or references) are time consuming and costly. Judges are notoriously reluctant to express adverse criticisms, particularly in written form. Also, common sense clearly indicates that when a man refers to three or four persons most competent to express opinions regarding his personal characteristics, he will name only those who will be reasonably sure to give him "good references." Finally, the difficulties inherent in check lists and rating scales set definite limitations when opinions are solicited by such means. And so, one may well ask, "Is it possible to obtain valid judgments concerning personal characteristics *from judges personally selected by a candidate?*"

II. THE PROBLEM

If judges named by a candidate are impartial observers and recorders of behavior, we should expect the mean of the distributions of the ratings of such judges to be lower than the mean of the distribution of the self-estimates of the students. In other words, unless we were dealing with a group of excessively modest persons we should expect the usual over-estimation tendency of self-raters to be in evidence when their scores were compared with the scores of their judges. The problem in this particular study was to compare the self-estimates of a group of college seniors and graduate students who reacted to a "Judgment Blank for Estimating Personal Traits" with the results obtained from judges designated by the students, who rated the students on the same list of traits.

III. GATHERING THE DATA

Two forms of a "Judgment Blank for Estimating Personal Traits" were prepared. Each form described twenty traits. All descriptiv phrases in Form A-1 were stated in the first person, singular. Form A-1 was used in obtaining the self-estimates of one hundred thirty-seven college seniors and graduate students at Cornell University.

4. Consider whether you will have the ability to "blaze new trails" in your profession. At present, can you see new relationships in old material? Do you have a questioning attitude toward things and people? Are you interested in finding solutions to problems? Are you original?

Can work out solution if problem is clearly stated for me.	Define problems clearly. Able to develop ideas independently.	Believe things as they are cannot be improved.	Have fanciful and impractical ideas which I insist upon putting into operation.	Disturbed by present conditions, but not able to work out solution.

8. Consider your mental balance. Are you stable? Do you "go to pieces" easily or "fly off the handle" readily? Are you inclined to be unduly cheerful and optomistic? Or, unduly depressed and morose? Are you taciturn and silent? Do you avoid company? Are you shy and retiring? Do you think you would have a "nervous breakdown" if things went wrong?

A moderate daydreamer. Sometimes depressed or discouraged, but not to an alarming extent.	Usually cheerful, but really not responsive to subtle changes in the environment.	Am in close touch with reality. Very stable. Mood nearly always appropriate to the occasion.	High strung; erratic. Often difficult to get along with.	Easily hurt or angered. Believe I am not given a fair chance. Blame others for my own mistakes.

14. How well do you conform to authority? Do you accept suggestions readily? Or, are you inclined to question what you hear and see? Are you too suggestible? Are you able to maintain a healthy balance between blind acceptance and thoroughgoing skepticism?

Conform most of the time.	Occasionally make emphatic objection.	Accept statements; am not critical.	Resent practically all authority.	Recognize only certain authorities.

FIG. 1.—Sample statements, Form A-1, cast in first person, singular. See Table I for percentages of self-raters checking each item.

Form A-2 dealt with the same list of traits, and was cast in the third person, singular. Form A-2 was used by three hundred fifty-three judges who rated the students. The forms were identical in content. They differed only in the point of reference. Form A-1 was designed to induce the mental set, "*My* personal characteristics are . . . " Form A-2 was designed to create the mental set, "*His* personal characteristics are . . . " Figure 1 reproduces three of the twenty parts of Form A-1. Figure 2 shows three parts of Form A-2. At first

glance, these forms appear like graphic rating scales. That, however, is not the case. In each instance, the descriptive phrases are arranged in random order in order to force the raters to read all descriptive phrases before indicating a judgment. The forms are essentially check lists of terms and phrases. They have been referred to as "scrambled graphic rating scales."

4. Consider whether he will have the ability to "blaze new trails" in his profession. At present, can he see new relationships in old material? Does he have a questioning attitude toward things and people? Is he interested in finding solutions to problems? Is he original?

Can work out solution if problem is clearly stated for him.	Defines problems clearly. Able to develop ideas independently.	Believes things as they are cannot be improved.	Has fanciful and impractical ideas which he insists upon putting into operation.	Disturbed by present conditions, but not able to work out solution.

8. Consider his mental balance. Is he stable? Does he "go to pieces" easily or "fly off the handle" readily? Is he inclined to be unduly cheerful and optomistic? Or, unduly depressed and morose? Is he taciturn and silent? Does he avoid company? Is he shy and retiring? Do you think he would have a "nervous breakdown" if things went wrong?

A moderate daydreamer. Sometimes depressed or discouraged, but not to an alarming extent.	Usually cheerful, but really not responsive to subtle changes in the environment	Is in close touch with reality. Very stable. Mood nearly always appropriate to the occasion.	High strung; erratic. Often difficult to get along with.	Easily hurt or angered. Believes he is not given a fair chance. Blames others for his own mistakes

10. Consider the tendency "to do things differently, to be able to modify habitual procedure, to depart from routine." How flexible is he? Can his opinions be changed after he has made up his mind? Can he adjust to new people? Can he fit into new situations?

Modifies methods with slight effort.	Quite rigidly bound by routine.	Slightly adaptable and adjustable.	Plastic, but within certain limits.	Seldom responds in the same way twice.

FIG. 2.—Sample statements, Form A-2, cast in third person, singular. See Table I for percentages of judges checking each item.

Cornell University graduate students in Education[1] and seniors in classes in Psychology and Personnel Administration in the Hotel Management Course, were invited to fill out Form A-1. Then three copies of Form A-2 were given to them, with the request that they select friends or acquaintances to fill them out. It was pointed out that it would be desirable if responsible business acquaintances were used as judges so that the records could be consulted later in answering requests for information when the student was being considered for a

[1] I am indebted to Professor Paul J. Kruse for assistance in obtaining the cooperation of graduate students for the purposes of this study.

position. In the majority of cases this suggestion was followed. The checked copies of Form A-2 were mailed to the college by the judges, and were not seen by the individual students.

IV. RESULTS

A tabulation was made, showing the frequency with which each of the one hundred descriptive phrases, arranged in groups of five each, under twenty main headings, were checked. Table I presents data showing the per cent of students and judges who checked each descriptive phrase.

Eight descriptive phrases were avoided consistently by all students who were estimating their own personal traits. No one of the one hundred thirty-seven who cooperated in this study described himself in the following terms.

Physically weak. Probably would break under strain of continued work. (1*A*)[1]
Dull, stupid, lifeless, mentally inert. (2*E*)
Rarely and grudgingly work with others. (3*D*)
Cooperate reluctantly and only when forced by circumstances. (3*E*)
Over-dressed. Arouse antagonism. Forward. (7*A*)
Am widely known to be untrustworthy and unreliable. (12*E*)
Satisfied with present. Live from day to day. (13*D*)
Antagonize others. Bully. Coerce my associates. (20*C*)

There were six statements that were checked once (but not by the same self-rater in each case.) They were:

Believe things as they are cannot be improved. (4*C*)
Easily hurt or angered. Believe I am not given a fair chance. Blame others for my own mistakes. (8*E*)
Am perfectly happy and contented as I am.
Have no desire to enlarge activities. (11*A*)
Do not believe in exerting myself.
Mentally and physically lazy. (16*A*)
Almost total lack of certainty. (17*E*)
Energy weak. Need urging to get started and constant supervision while executing the work. (18*D*)

There were eight descriptive phrases that were checked by more than fifty per cent of the self-raters. They were:

Make few minor errors. Am generally credited with good common sense. (5*A*)

[1] The Arabic numbers refer to the twenty traits which make up the scale. *A-B-C-D-E* refer to the order of appearance on the printed scale of the descriptive phrases, and not to the relative importance of the statements.

Dress for comfort. No outstanding characteristics. (7*C*)

Hold steadily but calmly to a definite goal. Make effort to attain it. (11*D*)

Can be relied upon to "deliver the goods" in so far as it is humanly possible. (12*D*)

Good worker, provided interest is sustained in job under way. (16*D*)

Admit possibility of being mistaken. (17*B*)

Energy concentrated in one direction. Show much initiative when dealing with things that interest me. (18*B*)

Become deeply absorbed in any problem that engages my interest. Not easily distracted. (19*B*)

Two descriptive phrases were not checked by a single judge as the traits of college seniors and graduate students were considered and indicated on Form A-2. They were:

Physically weak. Probably would break under strain of continued work. (1*A*)

Does not believe in exerting himself. Mentally and physically lazy. (16*A*)

Only one judge out of the three hundred fifty-three who filled out Form A-2 checked the following item in describing a graduate student, "Dull, stupid, lifeless, mentally inert." (2*E*) The following statements were checked by two judges.

Rarely and grudgingly works with others. (3*D*)

Over-dressed. Arouses antagonism. Forward. (7*A*)

Sometimes betrays confidence placed in him. Does not live up to expectations. (12*A*)

So slow in coming to decisions that he is of little value in most situations. (15*B*)

The following descriptive phrases were checked by more than fifty per cent of the judges.

Assistance eagerly sought by others and services cheerfully given. (3*C*)

Defines problems clearly. Able to develop ideas independently. (4*B*)

Makes few minor errors. Is generally credited with good common sense. (5*A*)

Dresses for comfort. No outstanding characteristics. (7*C*)

Is in close touch with reality. Very stable. Mood nearly always appropriate to the occasion. (8*C*)

Holds steadily but calmly to a definite goal. Makes effort to attain it. (11*D*)

Can be relied upon to "deliver the goods" in so far as it is humanly possible. (12*D*)

Seldom caught unprepared. Balances present and future needs. (13*C*)

Neither slow nor rapid; can't be hurried. (15*D*)

Excellent worker; requires little supervision. (16*E*)

Admits possibility of being mistaken. (17*B*)

Energy concentrated in one direction. Shows much initiative when dealing with things that interest him. (18*B*)

Becomes deeply absorbed in any problem that engages his interest. Not easily distracted. (19*B*)

Each of the college seniors and graduate students who filled out the self-rating blank (Form A-1), induced three judges to rate him on Form A-2. All rating blanks were scored by giving values of 5-4-3-2-1 to the descriptive statements which appear under the twenty main headings. These values ranged from five for the most favorable, to one for the least favorable statement. In this manner it was possible

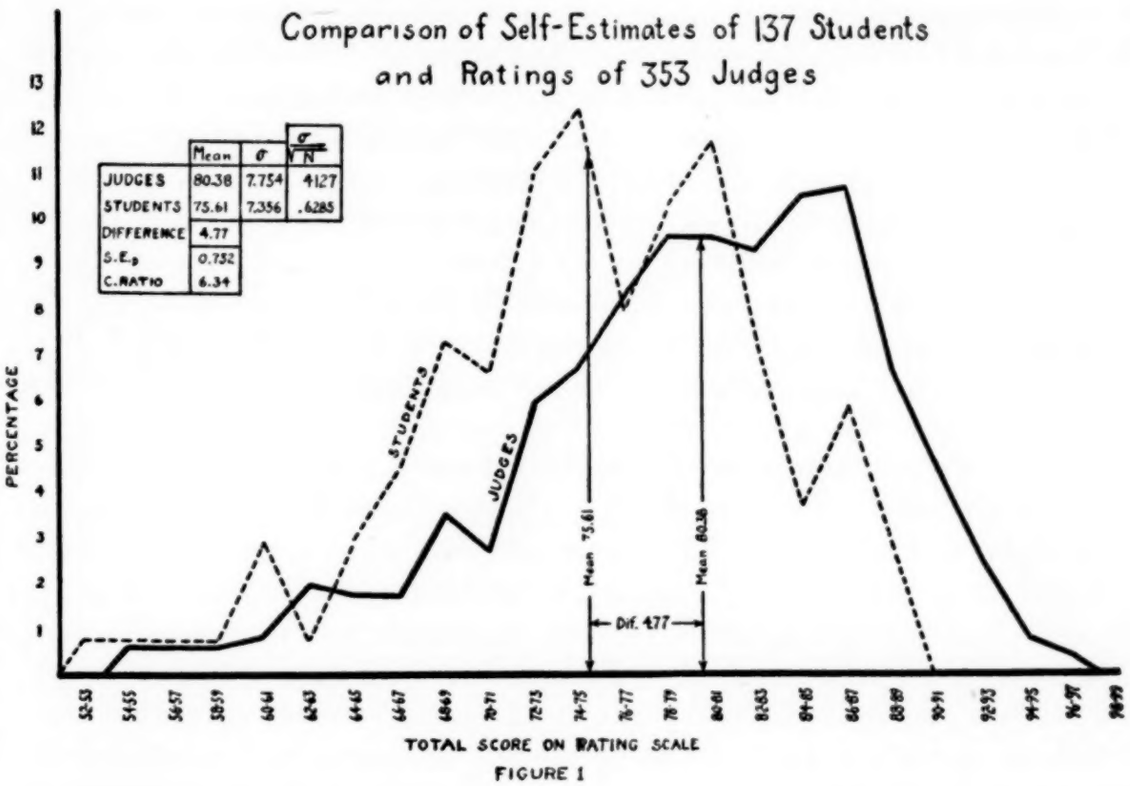

FIGURE 1

to score each completed rating blank. Theoretically, the ratings could range from twenty to one hundred. Actually, the range of ratings for Form A-1 was fifty-two to eighty-eight. The range for Form A-2 was fifty-five to ninety-seven.

The correlation, based upon one hundred four complete sets of ratings, between scores on Form A-1 (self-ratings) and Form A-2 (average of ratings of three judges) was $+.36 \pm .06$. The reliability of Form A-1 (split-half method) was $+.51 \pm .05$. The reliability of Form A-2, computed in the same manner, was $+.57 \pm .05$. The reliability of the entire scale, Form A-1, probably would be $+.68$. The estimated reliability of Form A-2 would be $+.73$.

The total scores for the one hundred thirty-seven Form A-1 records, and the three hundred fifty-three Form A-2 records were tabulated and plotted, in order to indicate the relationship that exists in these data between self-estimates and average ratings of judges. The mean of the distribution of the total scores of judges was 80.38. The difference between the means was 4.77. The standard error of the difference was 0.752. The critical ratio was 6.34.[1]

V. DISCUSSION

As this study of rating tendencies progressed it became apparent that more problems were raised than solved. The difficulty of preparing descriptive statements that were clear and concise was emphasized. Some adverse statements were avoided consistently by all self-raters and judges. Either the statements were so extreme as to represent degrees of traits that were considered non-existent in the group under consideration, or self-raters and judges were sensitive about checking statements of a decidedly negative character. On the other hand, thirteen favorable statements, out of a total of one hundred descriptive phrases, were checked by more than fifty per cent of the judges.

The judges were selected by the students, and were requested to complete and send in the blanks. The signing of the blanks, with the implication that they might be used in the future for placement purposes, may have induced a strong tendency toward lenient ratings.

Two other possibilities suggest themselves as we attempt to explain the marked differences in the distributions of the total rating scores of students and judges. Either the seniors and graduate students who cooperated in this study of rating tendencies represented an extremely modest group when appraising their personal traits, or, *they limited their requests for ratings to judges who were fairly certain to appraise them highly.* The latter hypothesis seems to deserve the greater weight. Ordinarily we assume that extreme rating tendencies may be corrected by averaging three or more judgments. This can hold true only when judges are selected by means of an impartial method.

When a student cooperates in an experiment of this kind and selects his own judges, or when an applicant for a position names three or four personal references, it appears likely that such judges and references will be named with a particular purpose in mind, *i.e.*, to

[1] I am indebted to Mr. M. W. Richardson for statistical aid in the preparation of this paper.

obtain as many favorable reactions as possible. Under such conditions it is obvious that judges are not chosen on an impartial basis. It is fallacious to assume that extreme rating tendencies will be corrected by averaging the responses of a number of judges so named.

A tendency exists for judgments obtained from persons named by the student to be surprisingly high. This does not mean that such judgments are without value. The relative position of an individual, in a group ranked by personally selected judges, may be of considerable importance. It appears, however, that the opinions of personally selected judges may be highly misleading, if we consider one individual at a time rather than a group of individuals. If we consider a single case we cannot say that he is judged better or worse than some other person. We may be influenced to consider him solely through the favorably prejudiced eyes of the judges he has named.

VI. SUMMARY

1. Two forms of a "Judgment Blank for Estimating Personal Traits" were prepared. Each form contained one hundred descriptive statements. Form A-1 (used for self-estimates) was cast in the first person singular, and was designed to induce the mental set, "*My* personal characteristics are . . . " Form A-2 (used by judges) was cast in the third person singular, and was designed to create the mental set, "*His* personal characteristics are . . . "

2. One hundred thirty-seven college seniors and graduate students rated themselves on Form A-1 of the "Judgment Blank for Estimating Personal Traits." Three hundred fifty-three judges (or references) cooperated by rating the one hundred thirty-seven students on Form A-2 of the scale.

3. Eight extremely uncomplimentary statements were avoided by all the self-raters. Eight extremely favorable statements were checked by more than fifty per cent of the self-raters.

4. The judges avoided two uncomplimentary statements in expressing their opinions. Thirteen extremely favorable statements were checked by more than fifty per cent of the judges who rated the students.

5. The reliability the of self-judgment blank, Form A-1, is +.68. The reliability of Form A-2, used by the judges, is +.73. The correlation between total test scores for Form A-1, and the average of the estimates of judges, obtained on Form A-2, is $+.36 \pm .06$.

TABLE I.—SHOWING PERCENTAGE OF STUDENTS AND JUDGES WHO CHECKED EACH DESCRIPTIVE PHRASE ON "JUDGMENT BLANKS FOR ESTIMATING PERSONAL TRAITS." IN EACH CASE, THE UPPER PERCENTAGE FIGURE INDICATES THE TENDENCY OF THE ONE HUNDRED THIRTY-SEVEN SELF-ESTIMATORS. THE LOWER PERCENTAGE FIGURE INDICATES THE TENDENCY OF THE THREE HUNDRED FIFTY-THREE JUDGES

Trait No.	Descriptive phrase, per cent					Total
	A	B	C	D	E	
1	0.0	2.8	30.8	47.5	18.9	100
	0.0	1.6	21.4	43.2	33.8	100
2	12.8	46.1	39.7	1.4	0.0	100
	9.8	30.6	48.1	11.2	0.3	100
3	23.3	37.3	39.4	0.0	0.0	100
	23.3	24.5	50.9	0.5	0.8	100
4	48.3	36.4	0.7	5.5	9.1	100
	38.5	55.3	1.3	0.8	4.1	100
5	55.6	1.4	9.9	19.7	13.4	100
	65.0	1.0	4.4	8.5	21.1	100
6	3.5	7.7	43.0	16.2	29.6	100
	2.9	4.8	33.3	15.9	43.1	100
7	0.0	18.3	60.6	19.7	1.4	100
	0.5	11.6	50.7	35.4	1.8	100
8	38.0	22.6	36.6	2.1	0.7	100
	10.8	24.1	62.5	1.8	0.8	100
9	14.3	15.7	10.0	20.0	40.0	100
	12.9	20.3	7.7	24.9	34.2	100
10	39.4	2.8	9.2	43.0	5.6	100
	37.1	5.6	13.0	41.1	3.2	100
11	0.7	6.4	5.7	83.7	3.5	100
	3.2	2.7	3.7	86.4	4.0	100
12	1.5	17.7	16.3	64.5	0.0	100
	0.5	4.4	14.4	79.7	1.0	100
13	27.1	4.3	44.3	0.0	24.3	100
	10.2	8.3	65.7	3.5	12.3	100
14	23.2	38.2	9.2	4.9	24.5	100
	26.9	36.2	12.4	4.4	20.1	100
15	4.2	2.1	10.6	38.0	45.1	100
	3.7	0.5	4.7	55.9	35.2	100
16	0.7	2.1	9.2	59.2	28.8	100
	0.0	1.3	7.1	34.1	57.5	100
17	9.1	65.0	21.7	3.5	0.7	100
	3.7	51.5	40.9	3.1	0.8	100
18	22.7	63.2	9.9	0.7	3.5	100
	12.5	61.2	21.5	1.6	3.2	100
19	19.6	54.5	2.1	4.2	19.6	100
	13.0	59.8	1.1	5.4	20.7	100
20	4.4	12.4	0.0	49.6	33.6	100
	8.8	12.4	6.3	45.0	27.5	100

6. The mean of the distribution of the total scores for the one hundred thirty-seven self-estimates was 75.61. The mean for the three hundred fifty-three estimates of judges was 80.38. The difference between the means of the distributions was 4.77. The standard error of the difference was 0.752; the critical ratio was 6.34. Since this difference is statistically significant, we may conclude that, in this group of personally selected judges, a pronounced tendency toward lenient ratings existed. This tendency may be considered all the more striking when we recall that our basis of comparison is the self-estimates of students who ordinarily are not inclined to under-estimate their own personal traits.

7. Evidence is presented in this study which indicates that personally selected judges tend to rate an individual higher than he rates himself. This may be due to a tendency on the part of the one naming judges, and requesting ratings, to limit his requests to persons who have previously given evidence of their good will, and who would be fairly certain to appraise him highly.

A STANDARDIZED TEST OF PROSE APPRECIATION FOR JUNIOR HIGH SCHOOL PUPILS[1]

HERBERT A. CARROLL

University of Minnesota

There appeared in the September issue of this magazine a paper which described and discussed a standardized test of prose appreciation for senior high school pupils. The writer has since developed a similar test for use in the junior high school. Since the fundamental assumptions and the test construction techniques are the same for the two forms, this article will be concerned only with such supplementary information as may be of value to the reader who wishes an explanation of the new data appearing in connection with the junior test. For all material common to the two tests, read the previously mentioned article.

VALIDITY

Sets I–VII and set XII, as they appear in the senior high school test, are used again in the junior high school form. Therefore, data on validity need be presented on the new sets only—numbers I and IX, following the junior high school order of arrangement. (See Table I.)

REMAINING DATA[2]

All further data pertaining to the junior high school but not to the senior high school test are presented in the following tables without comment.

SUMMARY

The Carroll Prose Appreciation Test for junior high school students is, like the senior high school test, based upon the assumption that the ability to appreciate literature can be measured by revealing the degree to which an individual discriminates among passages of varying worth. Eight of the selections used in the senior form have been incorporated in the junior form; two new ones have been added, the validity of these having been established by means of the two criteria: Source and a consensus of the ratings of sixty-five judges. The reliability of the junior high school form is .70 $\pm$ .019. The amount of time needed for taking the test is approximately thirty-

[1] Published by the Educational Test Bureau, Minneapolis, Minnesota. 1932.

[2] Data gathered with the assistance of Howard Jones, student at the University of Minnesota.

TABLE I.—VALIDITY ESTABLISHED BY TWO CRITERIA: SOURCE AND THE RATING OF SIXTY-FIVE JUDGES

	I					IX				
Selections	Critics	University instructors	H. S. teachers	Total	Source	Critics	University instructors	H. S. teachers	Total	Source
A—1	..	..	..	..	..	3	5	7	15	
2	..	..	..	..	..	5	10	9	24	2
3	..	1	3	4	..	6	4	8	18	
4	10	19	32	61	4	2	4	2	8	
B—1	..	1	1	2						
2	2	1	2	5	..	..	2	2	4	
3	8	17	29	54	3	2	3	1	6	
4	..	1	3	4	..	14	18	23	55	4
C—1	9	13	16	38	1	7	8	16	31	1
2	1	6	16	23	..	7	7	9	23	
3	..	1	3	4	..	1	7	1	9	
4	..	..	..	..	..	1	1	..	2	
D—1	1	7	18	26	..	5	10	3	18	
2	7	12	17	36	2	5	4	6	15	
3	2	1	..	3	..	6	9	16	31	3
4	..	..	..	..	..	..	..	1	1	
A	4	4	4	4	4	3	2	2	2	2
B	3	3	3	3	3	4	4	4	4	4
C	1	1	T	1	1	T	1	1	1	1
D	2	2	1	2	2	3	1	3	3	3

TABLE II.—RELIABILITY OF JUNIOR HIGH SCHOOL FORM OF THE CARROLL PROSE APPRECIATION TEST

N	*r*	PE of *r*
300	.70	.019

TABLE III.—ORDER OF DIFFICULTY OF THE TEN SETS OF SELECTIONS AS DETERMINED BY PERCENTAGE OF CORRECT SCORES MADE BY THREE HUNDRED JUNIOR HIGH SCHOOL PUPILS

Set No.....	I	II	III	IV	V	VI	VII	VIII	IX	X
Percentage of difficulty.....	61.7	50.5	49.6	44.0	43.4	39.9	39.8	37.9	32.1	30.7

TABLE IV.—DISTRIBUTION OF SCORES MADE BY JUNIOR HIGH SCHOOL PUPILS

Scores	Grade VII	Grade VIII	Grade IX	Total
52–55	0	1	1	2
48–51	1	2	1	4
44–47	1	7	17	25
40–43	7	13	31	51
36–39	8	17	37	62
32–35	20	30	64	114
28–31	34	56	65	155
24–27	56	67	57	180
20–23	75	74	61	210
16–19	88	68	29	185
12–15	78	41	26	145
8–11	28	18	9	55
4– 7	4	4	2	10
0– 3	0	2	0	2
N	400	400	400	1200

TABLE V.—MEASURES INDICATING THE DISTRIBUTION OF THE SCORES MADE BY PUPILS IN THE JUNIOR HIGH SCHOOL GRADES AND BY COLLEGE STUDENTS ON THE JUNIOR HIGH SCHOOL FORM OF THE CARROLL PROSE APPRECIATION TEST

Measures	High school pupils			College students
	Grade VII	Grade VIII	Grade IX	
N	400	400	400	100
H. P. S. (highest possible score)	60	60	60	60
Range	4–28	2–52	6–54	21–56
Mean (with probable error)	21.20 ± .258	24.36 ± .297	28.83 ± .306	39.52 ± .496
σ	7.64	8.81	9.08	7.36
Median (with probable error)	20.11 ± .322	23.62 ± .371	28.98 ± .382	40.00 ± .620
Q	5.22	5.89	6.47	5.47

TABLE VI.—DIFFERENCES WITH THEIR PROBABLE ERRORS, BETWEEN CERTAIN PAIRS OF MEANS

$$M_{(\text{College students})} - M_{(\text{Grade IX})} = 39.52 - 28.83 = 10.69 \pm .590$$
$$M_{(\text{Grade IX})} - M_{(\text{Grade VIII})} = 28.83 - 24.36 = 4.47 \pm .430$$
$$M_{(\text{Grade VIII})} - M_{(\text{Grade VII})} = 24.36 - 21.20 = 3.16 \pm .390$$

five minutes. Tentative norms have been established on twelve hundred representative pupils, four hundred from each of the grades VII, VIII, and IX

THE CONSTANCY OF THE IQ'S OF GIFTED CHILDREN

CLAUDE L. NEMZEK

University of Minnesota

From studies by Carroll and Hollingworth[2] and Lamson[4] the writer made a number of calculations concerning the constancy of the intelligence quotient.

Carroll and Hollingworth[2] reported the IQ's of fifty-two gifted children who were seven to nine years of age at the first testing and who had two IQ ratings from the Stanford-Binet and two from the Herring-Binet. The interval between the first and second testing was one year. The results are shown in Tables I and II.

Lamson[4] had fifty-three cases with two Stanford-Binet IQ ratings, forty-three of them having three. The interval between tests was one

TABLE I.—SOURCE OF DATA, TESTS USED, NUMBER OF CASES, CORRELATIONS BETWEEN TEST AND RETEST IQ'S, AVERAGE CHANGES IN IQ'S, RANGE OF THE MIDDLE FIFTY PER CENT OF CHANGES, SEMI-INTERQUARTILE RANGE OF CHANGES (Q), AND RANGE OF CHANGES

Data	Tests	N	r	Average change	Middle fifty per cent	Q	Range of changes
(1)*	H.-B.	52	.73 ± .044	9.06	−4.00 to 11.67	7.84	−19 to 22
(1)	S.-B.	52	.68 ± .051	9.02	−3.75 to 11.50	7.63	−21 to 32
(2)	S.-B. 1 and 2	53	.72 ± .045	7.85	−3.44 to 10.38	6.91	−21 to 25
(2)	S.-B. 1 and 3	43	.606 ± .065	8.95	−2.56 to 13.25	7.91	−18 to 31
(2)	S.-B. 2 and 3	43	.53 ± .074	9.37	−8.12 to 9.25	8.69	−21 to 33

* The numbers in this column refer to the references on pp. 609–610.

year; therefore, the interval between test one and test three was two years. The children studied by Lamson are, in some cases, the same subjects as Carroll and Hollingworth used. From the data presented by these authors, it is not possible to tell which children were used in both researches. The results derived from the data of Lamson are given in Tables I and II along with the findings from the study of Carroll and Hollingworth.

It may be of interest to compare the data presented in Tables I and II with some material concerning the constancy of the IQ's of

gifted children reported by Burks, Jensen, and Terman.[1] For seventy-two cases they found a mean IQ change of fourteen points, the SD of changes being 10.5. The interval between tests was six years. There was a distinct loss in mean IQ's on the retest. In the present study there was an increase in the mean IQ's of all retests. By the

TABLE II.—RESUMÉ OF THE NATURE OF THE DISTRIBUTIONS OF IQ'S

Data	Test used	*N*	Range of IQ's	Mean	SD	PE_{Mean}
(2)	H.-B. 1	52	108 to 174	135.8	13.8	1.3
(2)	H.-B. 2	52	116 to 173	138.8	13.3	1.24
			Difference	3.0	$\frac{D}{PE_{Diff.}}$	
			$PE_{Diff.}$	.935		3.21
(2)	S.-B. 1	52	133 to 190	152.3	13.85	1.3
(2)	S.-B. 2	52	128 to 188	157.9	13.45	1.26
			Difference	5.6	$\frac{D}{PE_{Diff.}}$	
			$PE_{Diff.}$	1.024		5.47
(4)	S.-B. 1	53	135 to 190	153.45	12.1	1.13
(4)	S.-B. 2	53	128 to 188	156.75	13.85	1.3
			Difference	3.3	$\frac{D}{PE_{Diff.}}$	
			$PE_{Diff.}$	.923		3.58
(4)	S.-B. 1	43	135 to 190	153.2	12.9	1.33
(4)	S.-B. 3	43	137 to 185	157.85	12.25	1.26
			Difference	4.65	$\frac{D}{PE_{Diff.}}$	
			$PE_{Diff.}$	1.15		4.04
(4)	S.-B. 2	43	128 to 188	156.9	13.4	1.38
(4)	S.-B. 3	43	137 to 185	157.85	12.25	1.26
			Difference	.95	$\frac{D}{PE_{Diff.}}$	
			$PE_{Diff.}$	1.29		.74

tetrachoric correlation procedure, Burks, Jensen, and Terman found *r*'s of .60 ± .10 for fifty-four cases and .65 ± .09 for seventy-three cases. The *r* for the group of fifty-four cases was found to be .81 and .77 by two other correlation methods.

In order to view the results obtained from the data of Carroll and Hollingworth and Lamson in light of characteristic studies based upon

unselected groups, various data from studies by Terman,[7] Rugg and Colloton,[5] Garrison,[3] and Rugg[6] are summarized in Table III. The facts indicate that the IQ is not as constant for the groups of gifted children as it is for unselected samples.

TABLE III.—RESULTS FROM SOME PREVIOUS STUDIES CONCERNING THE CONSTANCY OF THE IQ

Data	*N*	Interval between tests	Age range of cases	*r*	Mean IQ change	Range of middle fifty per cent of changes	*Q*
Terman,* 1919.....	428	1 day to 7 yrs.	3 to above 15 yrs.	.93 ± .004	4.5	−3.3 to 5.7	4.5
Rugg and Colloton, 1921	137	10 mos. to 1 yr. 4 mos.	6 yrs. to above 12	.84 ± .012	4.7	−2.3 to 5.6	3.95
Garrison, 1922.....	468	1 to 4 yrs.	.....	.88 ± .007	5.4	−2 to 4	3.0
Rugg, 1925.....	114	4 to 36 mos. median of 21 mos.	5 yrs. to 15 yrs., 11 mos.	.948 ± .006	3.1	−1.2 to 1.9	1.55

* Terman reported four hundred twenty-eight comparisons for three hundred fifteen cases.

CONCLUSIONS

1. The correlations between the repeated IQ ratings of these gifted children tend to be lower than the correlations reported by other investigators for more unselected groups.

2. The average change in IQ points, the range of the middle fifty per cent of changes, and the semi-interquartile range of changes tend to be higher than similar measures in previous studies based upon unselected samples.

3. In all cases the mean IQ on retesting was larger than the initial mean IQ, two of the differences being statistically significant. For these data there is a definite tendency for the IQ's of gifted children to increase. It may be justifiable to assume that the increases are primarily attributable to practice effects.

REFERENCES

1. Burks, B. S., D. W. Jensen and L. M. Terman: "Genetic Studies of Genius. III. The Promise of Youth." Stanford University Press, Palo Alto, 1930, 508 pp.
2. Carroll, Herbert A. and Leta S. Hollingworth: The Systematic Error of Herring-Binet in Rating Gifted Children. *Journal of Educational Psychology*, January, 1930, No. 1, Vol. XXI, pp. 1–12.

3. Garrison, S. C.: Additional Retests by Means of the Stanford Revision of the Binet-Simon Tests. *Journal of Educational Psychology*, 1922, Vol. XIII, pp. 307–312.
4. Lamson, Edna Emma: A Study of Young Gifted Children in Senior High School. *Teachers College Contributions to Education*, No. 424. Bureau of Publications, Teachers College, Columbia University, New York, 1930, 117 pp.
5. Rugg, H. O. and Colloton, C.: Constancy of the Stanford-Binet IQ as Shown by Retests. *Journal of Educational Psychology*, 1921, Vol. XII, pp. 315–322.
6. Rugg, L. S.: Retests and the Constancy of the IQ. *Journal of Educational Psychology*, May, 1925, Vol. XVI, pp. 341–343.
7. Terman, L. M.: "The Intelligence of School Children." Houghton Mifflin Co., Boston, 1919, 317 pp.

A BIOLOGICAL VIEW OF BEHAVIOR MODIFICATION

J. STANLEY GRAY

University of Pittsburgh

The problem of learning is the focal point of all psychology and education. It is the *sine qua non* of both fields. Psychology is valuable only to the degree that it explains how organic behavior is modified. Education is valuable only to the degree that it modifies behavior. And yet the problem of learning is perhaps the most befuddled issue in either field. Psychologists differ from each other in explanation and educators differ from each other in application or practice. Learning is a moot question to say the least.[1]

The writer believes that this problem belongs properly to the field of biology and only there can a scientific solution be found. Whether a biological explanation of organic learning should be applied to the human level or not is another question, but one which will gradually drop out with the progress of civilization. However, the biologists have little to say regarding the problem of learning as such. Instead they talk of changes in "protoplasmic composition" and in "organic pattern." While they largely ignore the problem of learning *per se*, they do give us a scientific explanation of the physical basis of learning. So perhaps the first problem which we should consider is that of the relationship of behavior to structure, or of behavior modification to structural modification.

I

The view that organic behavior is determined by the structure of the organism is supported by both psychologists and biologists. Meyer identifies behavior modification (learning) with changes in neural resistances.[2] Thorndike explains learning as being a process of "perpetuating, eliminating and modifying" "preformed bonds" or "original tendencies." He explains bonds to be "threads of specialized protoplasm each connecting one part of the body with another."[3] Smith and Gutherie say that "Conditioned responses (learning) involve

[1] See Bode, B. H.: "Conflicting Psychologies of Learning." Heath, 1929.

[2] Meyer, Max F.: "Psychology of the Other One." Missouri Book Co., 1922, p. 127

[3] Thorndike, E. L.: *Educational Psychology*, Vol. I. Columbia University, 1914.

the formation of new pathways in the neural system."[1] Other prominent psychologists who identify learning with morphological changes are Weiss,[2] Hunter,[3] Watson,[4] Dorsey,[5] and Sandiford.[6] With biologists, the view is even more prevalent. Herrick states that "If internal changes persist after the completion of reaction in such form as to alter the character of succeeding reactions, the protoplasm may be said to have been modified and these subsequent reactions which show the effect of this change may be classed as modifiable behavior."[7] Child explains that "Organismic behavior presupposes the existence of an organism, and is dependent upon the mechanisms present in that organism and their pattern of integration. The behavior of an *Amoeba* when exposed to white light of a given intensity is different from that of a starfish and this again from that of an earth worm. The behavior of a green plant under the same conditions differs widely from all of these. Such differences in behavior, whether of different individuals of the same species, of different related species, or of members of widely separate groups have their foundation in differences of some sort in the make-up of these different organisms."[8] The idea is likewise emphatically stated by Newman, "Structure is meaningless without function, and *vice versa*. In last analysis, function determines structure and structure in turn conditions function. The two are merely parts of the same thing, opposite sides of the coin."[9] Coghill states, "The mechanism we described (that of the Amblystoma) may therefore be regarded as the structural counterpart of the form of the behavior pattern."[10]

And so we might continue to state similar opinions of both psychologists and biologists. The fact that behavior is determined by structure seems to be pretty well established in the scientific world. Change the structure and you will change behavior—seems to be no

---

[1] Smith and Gutherie: "General Psychology in Terms of Behavior." Appleton, 1921.

[2] Weiss, A. P.: "Theoretical Basis of Human Behavior." R. G. Adams, 1929.

[3] Hunter, W. S.: "Human Behavior." University of Chicago, 1928.

[4] Watson, J. B.: "Psychology from the Standpoint of a Behaviorist." Lippincott, 1924.

[5] Dorsey, G. A.: "Why We Behave Like Human Beings." Harpers, 1925.

[6] Sandiford, Peter: *Educational Psychology*. Longmans, 1928.

[7] Herrick, C. J.: "Neurological Foundations of Animal Behavior." Holt, 1924, p. 283.

[8] Child, C. M.: "Psychological Foundations of Behavior." Holt, 1924, p. 3.

[9] Newman, H. H.: "Outlines of General Zoology." Macmillan, 1924, p. 54.

[10] Coghill, G. E.: "Anatomy and the Problem of Behavior." Cambridge University Press, 1929, p. 105.

less true of an organic than of an inorganic machine. Both represent the movement of complicated arrangements of lever arms, weights and fulcrums. The fact that the organic machine is constructed out of protoplasm, and that the manufacturing agents (genes) are located within the structure itself, is not evidence that it is non-mechanical or that it transcends the laws of physical causation.[1] The laws of the behavior of protoplasmic machines are different from the laws of the behavior of inorganic machines although neither violate the basic mechanical laws of lever movement. Inorganic machines of the same pattern but made out of different material do not behave alike. Organic machines differ from inorganic ones not only in material but also in pattern. That the resultant behavior is different can only be expected, but that that behavior is non-mechanical is an unscientific postulation. A prominent biologist has well concluded that while "the mechanistic or machine theory of living beings is not fully established, and it *may* not be adequate or even true; yet until every other possibility has really been exhausted, scientific biologists should hold fast to the working program that has created the science of biology. The vitalistic hypothesis may be held as a matter of faith; but we cannot call it science without misuse of the word."[2]

However, structure is not the only determining factor in organic behavior. " . . . environment is an essential factor, not only for structural existence, but also for the development of the form and structure and the physiological relation of parts characteristic of the individual."[3] "The organism as a pattern, as a mechanism, has no meaning except in relation to environment."[4] "Every movement is preceded or accompanied by another; energy is conserved; changes in the relationship of the structure and its environment elicit or cause its responses; behavior is not self-originated or spontaneous."[5] In other words, all behavior must be stimulated or caused, and the type of behavior depends both on the nature of the behaving mechanism and the nature of the cause. To understand behavior we must understand both the organic mechanism and its *milieu*. As Young points out,[6] any

[1] Gray, J. Stanley: "A Biological View of Original Nature." *Educational Administration and Supervision*, Dec., 1930.

[2] Newman, H. H.: *Op. Cit.*, 1924, p. 47.

[3] Child, C. M.: "The Individual and Environment from a Physiological Standpoint," in *The Child, The Clinic, and The Court*. New Republic Co., 1925.

[4] Child, C. M.: *Op. Cit.*, 1924.

[5] Rexroad, C. N.: "General Psychology." Macmillan, 1929, p. 31.

[6] *Social Psychology*. Crofts, 1930, p. 35*f*.

attempt to separate the two is a scientific fallacy of the greatest magnitude.

Without question, if we accept the conclusion of the scientists, we must recognize the organism as a physical entity whose behavior is controlled by the physical forces of nature. Organic behavior is no less *determined* than is movement of inorganic machines. It is always the direct resultant of a specific type of structure in a specific type of environment.

Now before we consider the problem of learning *per se* one more scientific conclusion should be noted. Protoplasmic modification (essential to learning) *is* behavior.[1] Metabolism is the basic process of all living material. Living protoplasm can do just one thing—metabolize or change. The nature of this change is determined, as mentioned above, by the *status quo* of the structure and the nature of the environment. In other words, the problem of learning is identical with the problem of behavior. Whenever there is behavior, there is structural change which determines, with the subsequent environment, subsequent behavior. As Child points out, the amount of structural change depends upon "the physicochemical constitution of each particular protoplasm."[2]

II

Our major problem can now be stated—What is the biological explanation of the alteration of organic structure? (This question is identical with the one—What is the biological explanation of learning?)

There are two phases of structural alteration, either of which will cause different behavior. First, a change of the organic substance itself will bring about a change in the behavior of it. Just what the chemical nature of the protoplasmic change is, has so far defied scientific investigation. In fact the chemical composition of protoplasm itself is largely unknown, but as Starling tells us "The gross methods of the chemist show him that the composition of the 'protoplasm' of the muscle cell is entirely different from that of a leucocyte or white

[1] Our conception of organic behavior must not be limited to merely the movement of somatic muscles. In fact, observable movement constitutes a very small proportion of organic behavior. As Herrick states, "The term 'behavior' usually refers to those visible movements of the body or its parts; but the invisible chemical, electrical and other changes are just as truly behavior of the living substance as is our running to catch a train." *The Thinking Machine*, University of Chicago, 1929, p. 70.

[2] Child, C. M.: *Op. Cit.*, 1924, p. 239.

blood corpuscle."[1] It is likewise evident that there is in protoplasmic change an assimilative or building up process (anabolism), and a dissimilative or tearing down process (katabolism). That these processes cause a change in the structural content of the protoplasm, even though the chemical nature of it remains a mystery, is the only fact which is significant for our purpose. Second, a change of behavior may be brought about by a change in the pattern or design of protoplasm. As Herrick explains, "All vital activities, all the things that living bodies do, are brought about by some changes in the arrangement of the stuff that these bodies are made of. The most significant of these changes are those chemical reactions between the different kinds of living substance, or protoplasm, that make up the organs and chemical reactions between this protoplasm and the oxygen, water, and food stuffs received from the outside."[2] Child sums up the importance of both phases of structural alteration as follows, "We may say that the general features of organismic pattern constitute the basis of physiological integration of the organism and so make it possible for it to react in one way or another as a whole, while the special mechanisms of reaction depend not merely upon the presence of an organismic pattern but upon the material, the kind of protoplasm in which the pattern exists."[3]

But the converse of this is likewise true. Structural changes not only cause behavior changes but behavior changes cause structural changes. Every organic movement leaves a structural after-effect which may or may not be permanent. Child says, "While it is perfectly true that the existing organismic pattern constitutes the general physiological basis for the behavior of the individual, it is also true that such behavior modifies the original pattern, determines new patterns within it or superimposes new patterns upon it. These changes may be either temporary or may persist. So far at least as the individual is concerned, behavior is the originator of new organismic patterns, new integrations, which are only potentially given in the original pattern."[4] Changes in behavior which are not due to structural changes are, of course, due to changes in environment.

Now what is the relation of behavior which brings about structural change to the behavior which results from structural change? A

[1] "Principles of Human Physiology" (Evans' revision). Lea and Febiger, 1930, p. 14.

[2] Herrick, C. J.: *Op. Cit.*, 1929, p. 70.

[3] Child, C. M.: *Op. Cit.*, 1924, p. 33.

[4] Child, C. M.: *Ibid.*, p. 72.

specific example will make this relationship more clear. Organism "A" (Johnny Brown) is so stimulated (asked to verbalize a section of a book) that it will behave in a way different from any way in which it has ever behaved before (repeat the multiplication tables). If this behavior is continued long enough (if the stimulation is continued) the organism will undergo certain permanent structural changes. These changes which have been brought about by a certain type of behavior will result in the same type of behavior but *with less stimulation necessary*. (The book is no longer necessary to stimulate the correct verbalization of the multiplication tables.) In other words, the organism learns to behave in a certain way by being stimulated to behave in that way. The amount of stimulation necessary to produce the behavior varies inversely with the amount of structural change which has been brought about by the preceding behavior. This simply means that a rat can learn to run a food maze by being stimulated to run a food maze. A dog can learn to perform tricks by being stimulated to perform tricks. A boy can learn to verbalize multiplication tables by being stimulated to verbalize them. The sort of behavior which brings about structural change is the sort of behavior which will result from that structural change.

We may now state specifically the biological position regarding learning. The organism learns to behave in a certain way when it is stimulated to behave in that way for a length of time sufficient for it to undergo those structural changes necessary to enable it to behave in the desired way when subjected only to initial stimulation. When an organism can complete a behavior series without error in response only to initial stimulation, it may be said to have learned the series. This is possible only after the organism has undergone a necessary amount of protoplasmic alteration. This alteration is physiological in character and takes place by a process known as metabolism. Fundamentally it is no different from structural alteration brought about by other causes, maturation for example. Learning is considered to be a distinct type of structural modification because of the distinction of cause rather than of process.

III

There are numerous conditions of the modifiability of organic structure, some of which deserve special attention. First, the parts of an organism vary in susceptibility to modification. It is quite evident that the metabolic gradient of the osseous system is less

modifiable than that of the neural system, for example. This is another way of saying that more learning is localized in the neural system than in the osseous system. As Herrick says, "All protoplasm can learn, but some combinations of living stuff can learn much more readily than others."[1] The protoplasmic structure of greatest susceptibility to modification is, of course, the cerebral cortex. Here modification may take place more rapidly than in any other part of the organism. However, this should not be interpreted as meaning that *all* learning takes place in the cerebral cortex. If we accept the organismal view of behavior,[2] we must conclude that no learning is completely localized in any part of the organism but that structural modification is body wide and alteration in any part affects the functioning of the entire organism. Consequently, the fact that some body parts are more modifiable than others has no practical significance.

Second, the susceptibility of an organism to modification varies from time to time. In general, modifiability decreases with development. Organisms in embryo form are more easily modified under experimental conditions than at any later time. For example, it has been found that in the early developmental stages of a frog embryo, polarity can be changed entirely and the head will develop from the opposite end of the embryo.[3] However, as the protoplasm becomes more highly specialized, its modifiability decreases. "Modifiability of the morphological pattern . . . decreases in general with the progress of development and of evolution."[4] There is much experimental work in the field of human education to substantiate this conclusion. Education, to be most effective, must take place early in the life of the child. This, however, does not suggest that the organism cannot be modified at later stages of development. So long as there is metabolism there is modifiability.

The physiological state of the organism varies from time to time and thus affects the rate of modifiability. Jennings pointed this fact out so early as 1906.[5] It can be verified by the naïve observation of any school teacher. An organism in a toxic physiological condition is not so modifiable as it is in a more normal state. Malnutrition is a

---

[1] *Op. cit.*, 1929, p. 231.

[2] Gray, J. Stanley: A Biological View of Organic Behavior. *Educational Administration and Supervision*, 1931.

[3] Rexroad, C. N.: "General Psychology." Macmillan, 1929, p. 73.

[4] Child, C. M.: 1924, p. 241.

[5] "The Behavior of the Lower Organisms." Columbia University Press, 1906, Chaps. X–XII.

frequent cause of abnormal physiological states which interfere with the process of school learning. The organism is most modifiable when its physiological condition is most normal.

Third, environment is vastly unstable and complex. It is impossible for an organism to be exposed to exactly the same sensory stimulation at two different times. The world about us is undergoing constant flux so that no two conditions can be identical. Not only this, but the organism has sense receptors located within itself (interoceptors) which could not possibly be stimulated at different times in the same way. Never does an organism respond to a single stimulus but always to complex stimuli situations. Hundreds of individual receptors are being affected constantly by hundreds of stimuli, both outside and inside of the organism. The particular ones which are dominant vary from time to time but never are the dominant stimuli independent of other stimuli. Never are some of the organism's receptors entirely functionless while others are functioning. The ears are still hearing even though visual stimuli may be dominant. Certainly it cannot be said that the auditory stimuli are entirely impotent. A particular stimulus or group of stimuli is always within a stimulus situation, some of which is located inside the organism.

The *gestalt* psychologists have been very effective in emphasizing the complexity of the stimulus situation. It is true that some of them have complicated it by attaching teleological explanation, but on the whole their contribution has been positive. They have used the term "pattern" to refer to the stimulus situation. Wheeler has clearly stated the *gestalt* position—" . . . any reaction of the human organism-as-a-whole is a unified response made to a total situation of some kind and, if to a specific detail, always to that detail in relation to other details. We may call this total situation a stimulus-pattern or arrangement of stimuli."[1]

IV

Now what are the implications of this interpretation of learning for education? It is quite evident that modern biology has added nothing essentially new to the common sense conception of learning. The conception that the organism learns to do by doing is by no means new in the history of human development. However, biology has added evidence and has made a naïve idea scientific. By implication, it has refuted those theories of learning which are based on non-scientific

[1] Wheeler, R. H.: "The Science of Psychology." Crowell, 1929, p. 77.

postulates. It has rescued the organism from the field of teleological phenomena and definitely established it as a part of a physical universe which operates according to the laws of physical causation (known to the scientists as the first law of thermo-dynamics or the law of the conservation of energy). Such terms as "insight," "experience," "feelings," "will power," "consciousness," "intuition," etc.—unless they are interpreted to be forms of behavior[1]—must not be used in the description of educational techniques and certainly not considered as educational goals. Educators must use scientific terminology if they are to be understood even by each other.

The biological view of learning if applied to education will eliminate much of the freak technique which is now cluttering our educational system and to which educators themselves are beginning to respond negatively. Millions of dollars of public money are being wasted each year by educational techniques which are based on fallacious conceptions of the child. If educators would accept the fact that the child learns to do by doing (as does any other organism), they could then direct their energies to investigations of how the child should behave and to the development of schools (environments) for stimulating him to behave in the desired way.

A further implication of the biological view of learning is that attention must be given to the entire learning situation. If learning stimuli always occur within a situation, a configuration, a *gestalt*, it certainly becomes important for the educator to control not only the learning stimuli but the rest of the situation as well. Every object within visual range, every sound, every odor, the nature of the preceding behavior (which includes home conditions)—in short, every phase of the situation within which learning takes place, determines in part the nature of the learning. To control learning, it is necessary to control every phase of the learning environment. The personality of the teacher, the pictures on the walls, the students surrounding him, the weather outside, the book he studies, the clothes he wears, etc., all affect a boy's learning of the multiplication tables or his solution of a problem in social behavior. It is the function of educators to determine by experimentation the best learning environment as well as the best things to be learned. What color should the walls of the study room be? Should boys and girls study in the same room? What sort of desks and seats are best for study? Should there be coöperative or

[1] Gray, J. Stanley: Modern Educational Theory and Objective Psychology. *Educational Method*, May, 1931.

individual study, and in what branches? Should the recitation teacher be the study supervisor? Should the study room be of the conventional formal type or a sort of a lounge room (but used for study only)? Should the subject be studied every day or every other day? These and many other questions must be solved by educators if they wish to control student learning.

Perhaps the most important implication of this position is that it places the responsibility for pupil learning directly on the educator. If the pupil does not learn it is because the educator does not understand his physical limitations and does not arrange the school environment to fit him. The physical (and therefore behavior) *status quo* of the pupil has already been determined by inheritance and previous environment. To change him, the educator must set his stage to fit the peculiarities of that particular pupil. The school must be fitted to the child if the child is to be fitted to conventional standards. If the child does not learn, *it is the fault only of the educator*. Of course it may be that the educator's ignorance (speaking collectively) is too great to enable him to solve all the problems of pupil learning which are under his direction, but it does not relieve him of the responsibility. Educators are employed to teach pupils (to stimulate pupils to learn) and if they do not do so they have failed, regardless of the cause. Educational progress will never come from evading responsibility.

Thus we have learned from biology that a human organism undergoes structural alteration as determined by the behavior which he is stimulated to perform. This structural alteration consequently determines subsequent behavior. The implication is that learning to behave in a certain way is brought about only by behaving in that particular way. Thus the school should stimulate the child to behave today in the way he should behave in the future. The problem of the most desirable human behavior is quite evidently outside the field of biology. The biologist cannot suggest how humans should behave, but after that is determined he is certain that they will learn to behave in the desired way only by behaving in that way.

THE CURVE OF FORGETTING: ITS STATISTICAL APPLICATION

R. J. WHERRY

Cumberland University, Lebanon, Tennessee

This article makes no pretense of great originality. The curve in question is that discovered by Ebbinghaus. The writer found, however, that his classes in elementary psychology were much more impressed by the law when it was reduced to the form of an equation and actual examples were worked out in class.

Ebbinghaus held that the ratio of what is retained to what is forgotten varies inversely with the logarithm of time. Most textbook writers quote Ebbinghaus to this effect; many admit that the law is probably universally applicable, and then let the matter drop, with perhaps a word of warning that the law refers to changes of rate rather than to absolute rate, which is true.

The purpose of this article is to set the law up in the form of an equation,[1] and explain its use as an instrument of prediction. This should make its meaning and application to the matter of retention much clearer to those readers who are deficient in mathematical training.

We note first that the equation deals with a ratio, the ratio of what is *retained* (which we shall call R) to what is *forgotten* (which we shall call F); thus, we have R/F for the left hand member of our equation. Next, we notice that this is directly proportional to the reciprocal of the logarithm of the *time* (which we shall call T); thus, for the right-hand member of our equation we have $\frac{1}{\log (T)}$. We must, however, take into account in our equation the fact that the equation applies to variable rather than to a constant rate of decrease, and so we must include a term in the right-hand member of the equation which we shall call "k" to take care of this factor. Our total equation would then appear in the form

$$\frac{R}{F} = \frac{k}{\log T}.$$

[1] This equation without explanation or application is given by Scripture: "The New Psychology." Scribners, 1901, in the form

$$\frac{b}{v} = \frac{k}{\log t(c)}.$$

One more step is necessary, however, before we can use the formula. As a matter of convenience and in order to avoid obtaining logarithms of zero (when $T = 1$), which would make the formula useless, the formula should be in the form

$$\frac{R}{F} = \frac{k}{\log \frac{T(10)}{T_1}},$$

where T_1 is the time elapsed before the first test for retention is given.

In order that one may better understand the use of the equation, the following examples are given. First, let us consider Ebbinghaus's own work. Table I gives his results. Thus, we see that at the time of

TABLE I.—PERCENTAGE OF NONSENSE SYLLABLES RETAINED

Percentage (R).....	50	36	33.3	23	20
Time elapsed.....	1 hour	8 hours	24 hours	6 days	30 days

the first test for retention (T_1) the following values were obtained: $T_1 = 1$ hour, $R = 50$ per cent, and, by subtraction, $F = 50$ per cent.

Our next step is to discover the value of "k." We do this by substituting the known values in our equation. Hence we have, when $T = T_1 = 1$ hour,

$$\frac{50}{50} = \frac{k}{\log \frac{1(10)}{1}} = k = 1,$$

and since "k" is a constant for any given set of data, our equation for this particular set of data becomes

$$\frac{R}{F} = \frac{1}{\log T(10)}.$$

Now if we substitute the other values of T, being careful to reduce all terms to hours (unit in which T_1 is expressed), we find what percentage according to the rule, will be obtained. Thus, when $T = 8$ hours, we have

$$\frac{R}{F} = \frac{1}{1.90309} = \frac{x}{100 - x} = \frac{34.4}{65.6}.$$

In the same manner we obtain the other values listed in Table II.

TABLE II.—PERCENTAGE OF NONSENSE SYLLABLES RETAINED

Time	1 hour	8 hours	24 hours	6 days	30 days
Actual percentage	50	36.0	33.3	23.0	20.0
Predicted percentage	50	34.4	29.6	24.0	20.6

These results are very close but this is not surprising since the law was evolved from this very data. The test of its validity must come from its application to other data secured by other investigators using other types of data. Such an experiment is that cited by Dashiell[1] from the work of Henderson on the retention of prose selections, the results of which are given in Table III.

TABLE III.—AVERAGE PERCENTAGE OF LOSS IN REPRODUCTIONS

Material	Of words		Of meanings	
Time	2 days	4 weeks	2 days	4 weeks
School and high school pupils	13.5	29.7	8.3	21.8
College students and graduates	34.0	49.1	23.7	35.2

Applying the formula, by letting two days equal T_1, the finding "k," and then using that value in the equation to predict the amount retained after four weeks, by each group for each class of material, we secure the results shown in Table IV.

TABLE IV.—PERCENTAGE OF LOSS IN REPRODUCTIONS AFTER FOUR WEEKS

Material	Of words		Of meanings	
Age group	Children	Adults	Children	Adults
Actual percentage	29.7	49.1	21.8	35.2
Predicted percentage	25.1	59.0	16.2	40.0

These agreements are again fairly close (probably less than the probable error of the group), and indicate that the formula evolved is perhaps of widespread applicability.

[1] "Fundamentals of Objective Psychology." New York: Houghton Mifflin Co., 1928, pp. 365.

SUMMARY

1. From a statement of Ebbinghaus's law of retention, its equation, in working form, is found to be

$$\frac{R}{F} = \frac{k}{\log \frac{T(10)}{T_1}}.$$

2. The equation is found to predict, on the basis of the percentage recalled on the first testing, the amount which is retained at later dates (within a reasonable limit of accuracy) as is shown by its application to the experimental findings of Ebbinghaus and Henderson.

3. Sufficient explanation and one example worked out in detail will permit the laws demonstration to elementary students in psychology. The author has found this to create a greater interest in the subject than the mere verbal statement of the law.

THE RÔLE OF AFFECTIVE TONE IN LEARNING TO SPELL

ARDEN FRANDSEN

University of Minnesota*

This study is an attempt to measure the influence of the pleasant and unpleasant connotations of words upon the efficiency with which children learn to spell them.

Several investigators have endeavored to determine the influence of feeling tone on memory and retention, but the results have been extremely contradictory and inconclusive. Meltzer[5] has recently summarized the studies in this field, beginning with Colgrave in 1898 and ending with Chaney and Lauer in 1929. Of the twenty-six experiments reviewed, sixteen indicated that pleasantness had a facilitating effect on memory and retention, while in ten, the results were negative with respect to this factor. Five further studies, [3,4,6,7,8] published subsequent to Meltzer's review, present data in favor of pleasantness as a positive factor. The problem, however, cannot be solved by a balancing of the positive and negative evidence. Even though a majority of experiments agree in their general conclusions, the results are not strictly comparable because the same factors were not controlled in every case.

The present study differs from previous studies in three ways: (1) A different type of learning function was used. (2) The difficulty, and degree of affective tone of the learning units were more precisely determined. (3) The methodological assumptions which, as Meltzer has pointed out[6] (p. 401), complicate the interpretations of many previous studies, are believed to have been avoided.

PROCEDURE

The evaluation of the influence of the pleasant and unpleasant connotations of words upon the efficiency with which children learn to spell them involved two steps: (1) A list of words was obtained, in which for each word the degree of affective tone was measured in terms of the reactions of the subjects used in the experiment. (2) These words, classified according to degree of pleasantness, were submitted to the subjects for practice under conditions wherein the

* This study was made in the Psychological Laboratory at the University of Utah. The writer wishes to express his appreciation to Dr. Dorothy B. Nyswander and to Dr. M. C. Barlow for many helpful suggestions and criticisms.

time and frequency of repetition, recency, vividness of presentation, and word difficulty were equal for both pleasant and unpleasant words. The subjects in the experiment were two fifth-grade classes of forty-two pupils each.

The affective connotations of the words used in the experiment were determined by a modification of the controlled association method. To the verbal directions: "Write down all the words you can think of that suggest pleasant things to *see*, things you like to see" (similarly for hear, taste, smell, touch, and do), the subjects responded with a total of four hundred twenty-five different words associated with pleasant experiences. To the same directions, except that "unpleasant" and "don't like" were substituted for "pleasant" and "like" a total of three hundred eight unpleasantly toned words were given. From these two lists of words, two hundred twenty-two pleasant and two hundred twenty-two unpleasant words were selected for a quantitative evaluation according to *degree* of pleasantness. This was necessary because a word which is definitely pleasant or unpleasant for one individual may have a different connotation for a second person. Furthermore, a measure of the extent or intensity of the affective tone associated with a word was needed.

Therefore, on the assumption that words could be ranked according to degree of pleasantness on the basis of the relative number of individuals who said they were pleasant or unpleasant, the words were submitted to the pupils for "voting" on this factor. Each pupil was supplied with a copy of the entire list of words, arranged in a random order, and instructed to indicate by writing a plus (+) or minus (−) symbol before each word, whether the word suggested something pleasant—something he liked, or something unpleasant—something he disliked. The list of words thus obtained ranged in degree of pleasantness from 0 per cent, words which everyone judged unpleasant, to 100 per cent, words which everyone considered pleasant. From this scale of values three groups of words were arbitrarily defined. Words judged pleasant by only 0 to 24 per cent of the pupils were classified as unpleasant; those between 25 and 75 per cent were considered indifferent; and those between 75 and 100 per cent were regarded as pleasant.

The consistency of this ranking procedure is indicated by the Pearson r of .98 between the series of percentage indices assigned by two fifth-grade classes to a list of one hundred twenty words selected,

in a manner to be explained later, for the learning experiment. A coefficient of self-correlation of .976 for the fifth-grade class used in the learning experiment, tested on the same list of words with an interval of three months between the tests, is evidence for the reliability of the method. These unusually high coefficients may be accounted for partially by the fact that many of the words in this list were judged to be definitely pleasant or unpleasant by a large percentage of the pupils. This, of course, tended to weight the extremes of the distributions, which consequently enhanced the values of the correlations over what they might otherwise have been. The interpretation placed on these coefficients, however, appears none the less valid because of this fact; the correlations merely indicate that for the words in this list, the pupils' judgments concerning affective tone were highly consistent.

Before taking up the problem of equating the lists for difficulty, the results on the ratio of pleasant to unpleasant words in the vocabularies of the subjects will be considered. It will be recalled that in the original controlled association test the number of pleasant words obtained exceeded the number of unpleasant words. The ratio of pleasant words to the unpleasant words was fifty-eight to forty-two. Equal amounts of time were given for writing each class of words. Notwithstanding the fact that the same number of pleasant and unpleasant words were chosen from these lists for the "voting," the ratio of pleasant to unpleasant resulting from this procedure was 60.8 to 39.2. (In this case words judged pleasant by more than 50 per cent of the subjects were considered pleasant and those judged pleasant by less than 50 per cent unpleasant.) This tendency for experiences judged pleasant to predominate in number over those considered unpleasant has also been noted by several previous investigators.[2,1,9,6,3,7] These writers have interpreted their results to imply that the previous experiences of their subjects have been, on the whole, more pleasant than unpleasant. There is also the possibility, however, that some experience factor (ideas of convention, etc.) serves to inhibit responding with words associated with unpleasantness.

In order to be reasonably certain that the lists of pleasant and unpleasant words used in the learning experiment were of equal difficulty, all the words were dictated to the pupils as a preliminary spelling test. Only fifty words of this test were dictated on a single day. On the basis of this initial test, each word was given an index of difficulty in terms of the percentage of pupils misspelling it. The percentage

of misspelling in the list was found to range from 0 to 100 per cent. Obviously, words which everyone could spell correctly at the outset would be useless for measuring learning. Therefore, from the total list of words, a final list of one hundred twenty words was chosen, ranging in percentage of misspelling from 40 to 100 per cent, and in degree of pleasantness from 0 to 100 per cent. These words were selected in pairs, approximately complementary with respect to degree of pleasantness and roughly matched according to difficulty. That is, for a word classified as pleasant, which was judged pleasant by 80 per cent of the pupils, a corresponding word, judged pleasant by about 20 per cent of the subjects, was found for the unpleasant list; similarly for other percentages. At the same time, the words of each pair were matched for difficulty. The average percentages of pleasantness for the lists composed in this manner were 87.8 and 10.9 for pleasant and unpleasant words, respectively. The corresponding averages according to percentages of misspelling were 63.9 and 66.5.

The one hundred twenty words were now sub-divided into twelve lists of ten words each, six lists being composed of pleasant words and six lists containing only unpleasant words. These lists constituted the spelling lessons for one class of forty-two fifty-grade pupils for a period of fifteen days, including three days occupied with testing. Every fifth day a test was given covering the forty words which had been practiced on the preceding four days. The usual method of drill was employed, but in order to insure less than perfect learning, the customary time was cut in half. Special care was taken to equalize the frequency of repetition, manner, and vividness of presentation of each word. The drill procedure was standardized so that exactly the same steps with the same amount of time were employed in the teaching of each word. Furthermore, at the end of each practice period the spelling papers and tests were collected by the teacher so that no pupil had access to copies of the words for study outside the regular periods. Manner and vividness of presentation were controlled by routinizing the procedure and by having the same teacher conduct all the drill exercises. The factor of recency was controlled by the systematic distribution of pleasant and unpleasant words in the practice periods and end tests.

RESULTS

The absolute amount of gain made on each word was computed by subtracting the percentage of misspellings on the end test from that on

the initial test. Such units would not yield satisfactory indices of gains because the differences in percentage of misspelling on the preliminary test would not be given proper consideration. For example, a word first missed by 95 per cent of the subjects in the preliminary test would, of course, serve as a learning problem for a larger percentage of the subjects than a word missed by only 40 per cent. Thus, although the entire group was subjected to the practice drills, actually the words on which learning took place were studied by groups varying in number. In order to get a figure indicative of the relative amount of gain made on each word, the percentage of gain (percentage of pupils who learned to spell the word during the training period) was divided by the percentage of misspelling on the preliminary test. The relative amounts of gain made on the pleasant, unpleasant, and indifferent words are presented in Table I.

TABLE I.—AMOUNTS OF RELATIVE GAINS ON PLEASANT, UNPLEASANT, AND INDIFFERENT WORDS COMPARED

No. of words	Qualitative classification	Range of pleasantness, per cent	Average percentage of gain, per cent	Differences in gains, per cent
47	A. Pleasant	76 to 100	43.8	
				A − B
14	B. Indifferent	51 to 75	53.9	−10.1
				B − C
7	C. Indifferent	25 to 50	43.6	+10.3
				C − D
52	D. Unpleasant	0 to 24	38.15	+5.4
				A − D
120	Average	0 to 100	42.5	P − U = 5.65

In the second column of Table I, the one hundred twenty words are classified into three "qualitatively" different groups according to affective tone. The basis for this classification was the arbitrary divisions made in the range of pleasantness as determined by the "voting" technique. The fourth column shows the average amount of relative gain made for each class of words. In the last column the differences in gains made on the different groups of words are shown. The gain for the line on which the difference is written was subtracted from the gain on the preceding line, except that in the bottom line the difference listed is for the most pleasant and most unpleasant words The first column gives the number of words in each class.

It is apparent from an inspection of the first column that the most significant comparison to be made is that between the gains made on the forty-seven most pleasant words and the fifty-two most unpleasant words. The difference between the means of these two distributions of gains is 5.65 per cent. The ratio of the obtained difference to the SD of the difference is 1.42. Statistically, the difference is well within the range of chance fluctuations. By referring to Table I again, it may be observed that, on the whole, greater gains were made on the indifferent words than on words more affectively toned in either direction. The number of words, however, is too small to make a significant comparison. Probably a more pertinent interpretation to put on these results is to say that they reveal the unreliability of the difference just considered, between the most pleasant and the most unpleasant words. For example, the difference in gains on the most pleasant words and the words next in order of pleasantness (Group *B*) is greater, in favor of the latter, than the difference between the most pleasant and the most unpleasant words. Again, the difference in gains between the most pleasant and the group of words second in order of unpleasantness (Group *C*) is practically zero. The lack of a significant relationship between the degree of pleasantness and gains made in learning to spell the words is perhaps indicated more precisely by the correlation coefficient of .109 $\pm$ .06 between these variables. Since both reliability indices used to describe these data are in part functions of the size of the populations, it is impossible to decide with certainty whether the statistically insignificant $D/\sigma d$ and the coefficient of .109, which is only 1.8 times its PE, should be interpreted as indicating no, or slight, association. The precise degree of relationship can be determined only when it is certain that the sample of words measured is statistically adequate.

From this analysis of the data it appears that the relative amounts of gains made in learning to spell these words were influenced very little if any by the differences in affective tone of the words. There is a possibility, however, that the affective tone which was demonstrated for the words in the association test did not carry over to the spelling situation. In the association test the pupils were directed to respond directly to the pleasantness or unpleasantness associated with the words, whereas in the spelling situation the response was to a different aspect of the words. Because of this change in "reaction set" it may be that the affective tone of the words was not reacted to at all in the spelling exercises. However, the volunteered report of one child

during a spelling lesson, "One day you make me sad and the next day you make me glad," offers some evidence, very slight to be sure, that the affective aspect was a part of the situation. This is only a suggestion, of course, and needs to be investigated further. In any case, however, the conclusion is true for the conditions under which the experiment was conducted; and these conditions were typical of the spelling class situation in the school.

CONCLUSIONS

1. Of the words given by the subjects as responses to the "controlled association test," there was a preponderance of words associated with pleasant experiences.

2. In a typical school drill situation, the efficiency with which children learned to spell words was shown to have very little or no relation ($r = .109 \pm .06$) to the differences in the pleasantness or unpleasantness of the words.

REFERENCES

1. Flugel, J. C.: A Quantitative Study of Feeling and Emotion in Everyday Life. *British Journal of Psychology*, Vol. XV, 1924–1925, pp. 318–355.
2. Henderson, E. N.: Do We Forget the Disagreeable. *Journal of Philosophy, Psychology & Scientific Method*, Vol. VIII, 1911, pp. 432–438.
3. Jersild, Arthur: Memory for the Pleasant as Compared with the Unpleasant. *Journal of Experimental Psychology*, Vol. XIV, 1931, pp. 284–288.
4. Koch, H. L.: The Influence of Some Affective Factors upon Recall. *Journal of General Psychology*, Vol. IV, 1930, pp. 171–190.
5. Meltzer, H.: The Present Status of Experimental Studies on the Relationship of Feeling and Memory. *Psychological Review*, Vol. XXXVII, 1930, pp. 124–139.
6. Meltzer, H.: Individual Differences in Forgetting Pleasant and Unpleasant Experiences. *Journal of Educational Psychology*, Vol. XXI, 1930, pp. 399–409.
7. Stagner, Ross: The Redintegration of Pleasant and Unpleasant Experiences. *American Journal of Psychology*, Vol. XLIII, 1931, pp. 463–468.
8. Thomson, Ruth H.: An Experimental Study of Memory as Influenced by Feeling Tone. *Journal of Experimental Psychology*, Vol. XIII, 1930, pp. 462–468.
9. Wohlgemuth, A.: The Influence of Feeling on Memory. *British Journal of Psychology*, Vol. XIII, 1923, pp. 405–416.

THE RELATIVE EFFICIENCY OF THE WHOLE, PART AND PROGRESSIVE PART METHODS WHEN TRIALS ARE MASSED—A MINOR EXPERIMENT

J. B. STROUD AND C. W. RIDGEWAY

Kansas State Teachers College

INTRODUCTION

In his work on methods of learning, Pechstein found that the superiority of the whole method over the part method of learning obtained only in case the trials were distributed.[1] On the contrary, he found that the part method was superior to the whole method when practices were massed. This condition held true for stylus maze and nonsense syllable material. In the maze studies he found a variation of the part method (the progressive part method) to be more economical than the whole method with practices massed or distributed, but slightly less economical than the part method with practices massed. In his study with nonsense syllables, he found the progressive part method to be generally superior to either the whole or part method when practices were massed.

The aim of the present experiment is to carry out certain features of Pechstein's experiments, namely, the comparative efficiency of the whole, part and progressive part methods when practices are massed with meaningful material. Inasmuch as circumstances frequently make distribution of practices inadvisable, it seems desirable to determine whether the part and the progressive part methods are generally superior to the whole method when practices are massed.

PROCEDURE

The material used in this experiment was poetry. It consisted of five stanzas, four lines to the stanza. Three such selections were used, all taken from the same author and essentially the same type of poetry both as to subject matter and metrical arrangement.

The subjects were thirty-six college students. Each subject learned each of the three poems, one poem by the whole method, another by the part and the other by the progressive part method. The poems proved to be practically equal in difficulty. Nevertheless,

[1] Pechstein, L. A.: Whole *vs.* Part Methods in Motor Learning. *Psychology Monographs*, Vol. XXIII, No. 2, 1917.

Whole *vs.* Part Methods of Learning Nonsensical Syllables. *Journal of Educational Psychology*, Vol. IX, 1918, pp. 381.

a counterbalanced order was used. Practice effect was controlled by counterbalancing the order with respect to the methods.

In all cases the material was read aloud at a rate suitable to the subject. The subjects were instructed to read a stanza or stanzas through, depending upon the method used, without re-reading any special parts. The material was covered up immediately after each reading and the subject reproduced as much as possible of what he had just read. Reading and recall alternated until the poem could be reproduced without error. The subjects learned one poem a day for three consecutive days.

With the whole method the entire five stanzas were read as a unit. With the part method each stanza was learned separately to the point of one correct repetition, followed by an integration of the selection as a whole. In the case of the progressive part method the subjects learned the first and second stanzas independently and then read them as a unit of two stanzas until they could repeat them in this manner. The third stanza was then learned independently. Whereupon, the first three stanzas were integrated. This process was continued until the entire selection could be repeated without error.

RESULTS

The table below gives the average time and the average number of trials required to learn the poems by the three methods.

TABLE I.—THE RELATIVE EFFICIENCY OF THE WHOLE, PART AND PROGRESSIVE PART METHODS

Criterion	Whole	Part	Progressive part
Time	31′37″ ± 1.12	29′17″ ± 1.22	27′21″ ± 1.14
Trials	14.50 ± 0.49	10.98 ± 0.47	10.50 ± 0.48

In computing trials for the part and progressive part methods, each reading of a stanza, together with the recall, was considered one fifth of a trial.

In these results the whole method is at a disadvantage when compared with either of the others in terms of both criteria. With time as a criterion the differences are not statistically reliable. The difference between the average time for the whole method and the progressive part method is 2.6 times the PE of the difference. The PE's of the differences between the other means are practically as large as the difference. When the three methods are compared on the basis of trials, the part and progressive part methods are significantly superior to the whole method.

The fact that the differences between the means of the time scores are not statistically reliable may be due, in part, to the high degree of variability of such scores. The standard deviations of the time scores are two and three times as great as those of the trial scores. Owing to the fact that subjects differ so widely in rate of reading and in time consumed in recall, trial scores may be considered more reliable than time scores when time intervals of presentation and recall are uncontrolled.

These results tend to show that the same conditions regarding the superiority of the part method of learning maze and nonsense material, when practices are massed, also hold true for meaningful material. These data may also be regarded as contributing to the general proposition that the part and progressive part methods are superior to the whole method when practices are undistributed.

Incidentally, complete relearning scores were obtained on eighteen of the subjects. Relearning was accomplished one week after learning. All subjects relearned by the whole method. It seemed that this method would give the most reliable measure of retention. It was probable that some subjects would relearn the poems in one or two trials. Such subjects would be at a disadvantage relearning by the part and progressive part methods.

TABLE II.—THE RELATIVE EFFICIENCY OF THE WHOLE, PART AND PROGRESSIVE PART METHODS IN TERMS OF RETENTION

Criterion	Whole	Part	Progressive part
Time	6′26″	6′26″	7′11″
Trials	3.0	2.70	3.10

From the standpoint of retention, there seems to be little to choose between the three methods.

CONCLUSIONS

1. The foregoing results indicate that the whole method is less economical than either the part or progressive part methods of learning meaningful material when practices are massed.
   (*a*) The differences in terms of trials are statistically reliable.
   (*b*) The differences in terms of time lack in reliability.
2. The relearning scores indicate that there is little difference between the three methods with respect to retention.

BOOK REVIEWS

LEE EDWARD TRAVIS. *Speech Pathology: A Dynamic Neurological Treatment of Normal Speech and Speech Deviations.* Pp. XXXIV + 331. New York: D. Appleton and Co., 1931.

Authors of books on speech disorders have not kept pace with research in the biological sciences. By adhering to old problems and even old data and neglecting materials of theoretical, clinical and scientific interest they have devitalized the field. Such, in brief, is the contention of Dr. Travis expressed in the preface of what he purports to be a comprehensive text of causes, diagnosis and treatment of speech disorders.

The chief source of stimulation for his action-current research as well as other studies which determined the dominant points of emphasis manifested by the author in this book is Dr. S. T. Orton. It was Dr. Orton who first familiarized him with the theory of bilateral rivalry which plays a significant part in Dr. Travis' work on stuttering. Related concepts that are made use of in this book are the contributions of Child, Herrick and also the Gestaltists.

The book starts with a chapter on the neuromuscular basis of speech and is followed by three short chapters on classification of speech disorders, causes, and examination methods. About one third of the book is devoted to a consideration of stuttering. This is done under the caption of "Disorders of Rhythm in Verbal Expression" in Chapter V. In Chapter VI are considered "Disorders of Articulation and Phonation" and in Chapter VII, "Disorders of Symbolic Formulation and Expression (Aphasia and Kindred Disorders.)" These topics are summarized in Chapter VIII. A lengthy appendix includes a list of stimulus-syllables, words and sentences for speech training and an index. A ten page glossary is included in the front portion of the book. The contents is outlined in all details. There are seventy illustrations in the book. The style of writing as well as organization is text-bookish.

The general attitude as well as the specific points of emphasis in this book are suggested by the following quotation from the chapter on stuttering:

"The neurological basis of stuttering, which is a lack of a sufficiently dominant gradient of excitation in the central nervous system, is largely a pathophysiological variant. We may think of its occurrence in terms of a margin of dominance which ranges in different individuals

from zero amount to a very large and safe amount. If the margin is small, stimuli of slight emotional value and conditions which bring about small amounts of physical and mental fatigue will produce stuttering. If there is equal dominance, we have the stutterer who stutters under all conditions of speech. If the margin is large, we have the normal speaker under practically all physical and mental conditions . . . we must not turn to stuttering from the consideration of social morbidity, inferiority feelings, oral eroticism, and the like without looking at what lies between them."

Because he considers symptoms of stuttering as mainly signs of the rivalry between the two sides of the brain in his treatment the possibility of reverting to left handedness is always considered and the carrying out of exercises to increase the dominance of one hemisphere is included.

Mental hygienists who consider stuttering as a personality disorder as well as students of stuttering like Dr. Fletcher who advocate environmental therapy for its treatment will not find the neurological theorizing of Dr. Travis adequate. Neither will they find in the case studies reported convincing evidence of the validity or reliability of the therapy which correlates with this theorizing. But the book is, nevertheless, a significant contribution to the field. Significant because it is the most exhaustive treatment of the subject from one important viewpoint and not because it is a well balanced, comprehensive treatment of the field of speech pathology. H. MELTZER.

Psychiatric Child Guidance Clinic, St. Louis, Missouri.

CARL MURCHISON, Editor. *A Handbook of Child Psychology.* Pp. 711. Worchester, Mass.: Clark University Press, 1931.

This handbook, the latest volume in the International University Series, is made up of twenty-two articles contributed, by eminent psychologists. In general, each author has summarized the present status in the field under discussion and indicated the problems being studied together with needed research. Among them are Gesell, Goodenough, L. S. Hollingworth, Kimmins, Margaret Mead, Piaget, Pintner, Terman and Woolley. A few contributors added materially to the usefulness of the text as a handbook with summaries at the conclusion of their chapters. Excellent bibliographies accompany each contribution.

The handbook is a reference which will hardly take the place of an introduction to child study. It will be an excellent supplement to such

a course and the more serious student possibly seeking to specialize in some phase of child psychology should not only be able, by reference to this text, to orient himself more readily as to the field most interesting to him but also find the directions in which research is most needed. The supervisor and administrator away from the graduate school for some time and pressed for time to keep up with the rapid specialization in child psychology should find the handbook and authoritative reference worth owning. J. H. COLEMAN.

Huntington, N. Y.

H. E. BUCHHOLZ. *Fads and Fallacies in Present Day Education.* Pp. XIV + 192. New York: Macmillan Company, 1931.

Here is a book that teachers and administrators should read, if for nothing more, than to observe certain phases of our profession as others outside the teaching profession see them. *Fads and Fallacies in Education*—the very chapter titles of the book reveal the point of view of the author: I. Leadership Pianissimo—the Tale of a Peanut; II. Ichabod Wants a Portfolio—Hence a Department of Education; III. More Money for Less Education—But America Is Young and Rich; V. The Poor Enriched Curriculum—a Triumph for the Progressives; VI. Prussianized Pedagogues—Teachers of Liberty; VII. Public Education—Fallacy of Equal Salary Schedules; and VIII. Sample School Textbooks—a Study in Polite Graft.

In the preface the author implies that his indictment of public education was so severe that an editorial conference led to the suggestion that someone write a foreword to provide all the "sugar and salt" for such a book.

Dr. W. C. Bagley has therefore written the Foreword. In it he suggests the dangers of satire—exaggeration. Thus his few pages endeavor to warn the readers that school people, as a group, are not a bad lot and that isolated examples of mistakes of teachers and school leaders do not necessarily mean that everything is wrong with education.

The reviewer feels that Mr. Buchholz has, in many ways, continued effectively his service as a "lay" critic of educational theory and practices, notwithstanding the result about which Dr. Bagley warns the reader—exaggeration. The reviewer would subscribe to Mr. Buchholz's arraignment of "more money for less education"—implied extravagance in tremendously increased expenses in education in recent years—and to "sample textbooks—a study in polite graft."

However one-sided they seemingly are the reviewer would even admit the force of the writer's charges against a national department of education. However, to imply that the typical college professor of education is his Mr. Horatio Bump "a peanut," is extreme. Moreover, for the purpose of substantiating his satire of progressive education the author cites a few examples—out of context—from two out of scores of progressive courses of study. In the reviewer's judgment this is absurd. Finally, the reviewer wonders if the writer is not at times guilty of the very thing he claims much modern education rests upon: Fallacious theories. For example, there is doubtless extravagance in the use of money for school buildings, equipment, etc., but seemingly the writer has ignored the inevitable consequences of a tremendously expanded program of democratic education, particularly on the secondary and college level, where specialization is rampant and this expansion comes largely from the pressure of democracy.

Yet, such books as this, in times such as these, should compel school people to "take stock." With all education under attack, real reconstruction and efficiency and economy can alone come from a trained profession. It is our obligation; it cannot be dodged, and it must be solved.

EARLE U. RUGG.

State Teachers College, Greeley, Colo.

CLAUDE C. CRAWFORD. *Studying the Major Subjects.* Pp. XIV + 384 Los Angeles: C. C. Crawford, 1930.

This book is one of a methods-of-study series written by the same author, sometimes in collaboration with others. The major subjects are: Literature, composition, foreign languages, mathematics, physics and chemistry, biology, history, other social studies, the practical arts, physical education and health, the fine arts. To each of these fields is assigned a chapter.

Dr. Crawford states that "the key word of the book has been HOW." No one will dispute this. Each one of the one hundred forty-two main sections, into which the chapters are divided, starts out with "How": "How to use the microscope," "How to learn the symbols for chemical elements," "How to avoid embarrassment in discussing topics relating to sex and reproduction," "How to see parallels between lower and higher forms of life" (a marvellous example of diplomacy; the word *evolution* is never specifically mentioned),

"How to choose a vocation," "How to develop a high moral character." The latter subject is disposed of in four pages.

Every one of these "How to do something or other" sections is subdivided into paragraphs. The topic sentence, written in italics, invariably appears first and is followed by illustrative material. Occasionally the suggestions are excellent. The student is advised to "check for major errors by means of estimates and approximations," and to "solve problems involving formulas" if he has difficulty in understanding the latter. He is told to write his report "as for one who did not see or know about the experiment." The author also considers it necessary to urge students to "read the sport page before and after the game," and to "run for office in student groups."

At the end of each of the "How" sections are assignments. The varied nature of these may be inferred from the following: "List all the different kinds of emotional expressions which at some time or another are suggested by your reading." "Tell what was your greatest difficulty in learning to pronounce a foreign language and how you overcame it." "Arrange a page in your note book in which to list biological terms whose meanings you wish to master." "Make a list of the qualities or elements which must enter into any useful article in order that it may be beautiful." "What were your principal difficulties in learning to dance?" "What can you do to beautify your own classroom and school environment?"

The material varies, but the form of the book is unyielding. No text of the Scholastics was ever written to a more rigid outline. The "How" sections, the topic sentences in italics, and the assignments recur in fatal monotony, and style is sacrificed to system.

The moral tone of the book is unimpeachable. The author is a believer in sublimation. "The reading of a beautiful love story is a splendid way to give expression to the mating instincts." "Habits of profanity and foul language are overcome by the cultivation of an interest in poetry and good literature. Similarly an interest in athletics may be a fine method of overcoming the temptation to use tobacco." The student is further urged to "do a real favor to someone and see how good it makes you feel," and as an assignment he is asked: "What can your class or school do to raise its moral tone?"

The book appears to the reviewer to be written with the high school pupil chiefly in mind. The first chapters seem more successful than the latter ones.

MELVIN RIGG.

Kenyon College.

LAWSON G. LOWREY (editor). *Institute of Child Guidance Studies.* Pp. 290. New York: The Commonwealth Fund, 1931.

The first volume of the studies from the Institute for Child Guidance includes sixteen contributions by ten contributors. Four of the contributors are social workers, four are psychiatrists and two are psychologists. Five of the studies included are from the pen of the editor, Dr. Lowrey, two are by Dr. H. M. Tiebout and two by Dr. Curt Rosenow. Six of the contributions are classified in the social field; six in the medical field; two as psychological and two as child training.

The two studies classified as contributions from the psychological field are "One More Definition of Heredity and Instinct" by Rosenow and "The Rorschach Test and Personality Diagnosis: The Feebleminded" by S. J. Beck. Only four of the studies included are at all quantitative in nature. These are; "Fingersucking and Accessory Movements in Early Childhood" by Dr. David Levy; "The Incidence of First-born among Problem Children" by Dr. Rosenow; "Restlessness in Infancy" by Dr. Hyman S. Lippman and the aforementioned study on the Rorschach test by Beck. Other papers which will be of interest to some educational psychologists are Dr. Tiebout's on stealing and some of Dr. Lowrey's. Most of the articles of interest to readers of this journal have been previously published in psychiatric and psychological journals.

A reading of these studies will give one a fair idea of the basic thinking underlying the mental hygiene approach to the study and treatment of child behavior as sponsored by the Institute of Child Guidance.

H. MELTZER.

Psychiatric Child Guidance Clinic, St. Louis, Missouri.

ICAL ROOM
RARY